

NASA Contractor Report 180780

Engineering Planetary Lasers For Interstellar Communication

(NASA-CR-180780) ENGINEERING PLANETARY
LASERS FOR INTERSTELLAR COMMUNICATION M.S.
Thesis (Maryland Univ.) 623 p CSCL 22B

N88-23825

Unclas
G3/18 0146049

Brent Sherwood

Contract NGT 21-002-823

May 1988

NASA

NASA Contractor Report 180780

Engineering Planetary Lasers For Interstellar Communication

Brent Sherwood
Department of Aerospace Engineering
University of Maryland
College Park, Maryland

Prepared for
Goddard Space Flight Center
under Contract NGT 21-002-823



National Aeronautics and
Space Administration

Graphic and Publication Services Branch
Goddard Space Flight Center

1988

PREFACE

This thesis for the degree of Master of Science in Aerospace Engineering is an order of magnitude bigger than most. Inquiring minds will want to know why, and hopeful readers deserve to know in advance. I did not start out intending to write a book. But even stripped to a skeleton of the work that supported it, this document encapsulates three years of NASA-sponsored, wide-ranging thought about the problem of interstellar communication: its purpose, its tools, and its implications.

The primary intent of the thesis is to demonstrate some mastery of the activity of design, as applied to advanced space systems. Practically absent from graduate engineering curricula traditionally thick with analysis, design demands a complementary sensibility. In addition to embracing the terror of evolving something real out of, tabula rasa, nothing, designers must reconcile an exhilarating array of conflicting attractors to solve any problem. Their goal is the intrinsic beauty of a workable, elegant solution; their creed is an even treatment of all pertinent facets.

Responsive and defensible thoroughness in that treatment means avoiding procrustean rigidity in laying bare the design alternatives, and rationales used to choose among them. For an unprecedented, speculative space mission, viable alternatives inevitably arise from research frontiers. I have

thus used this project to survey almost the gamut of advanced technologies applicable to space, following NASA's charter to transfer new technology through use and exposure.

Because many of those fascinating technologies, and other principles central to this work, will be unfamiliar in detail to most readers, I have devoted much space (entire chapters at times) to tutorial reviews. Not intended as exhaustive, the sections on cosmo-ethology, laser physics, phase control, communication theory, ring lasers, gravitational planetology, light diffraction, control technology, nuclear power, spacecraft subsystems and nanotechnology, and other briefer explanations throughout the text, are included to make the thesis as self-contained as patience permits.

Finally, because the broad range of topics covered could easily weave an impenetrable tangle of detail, I have favored what I hope will be an engaging literary style, to give the tenacious reader every chance of apprehending the fascinating complexity of designing planetary lasers for interstellar communication.

This work is dedicated to both Bobs.

ACKNOWLEDGMENT

Any project as complex as this must depend on the ideas, criticism and support of people "too numerous to list". I gratefully list them here nonetheless. The generosity and interest were theirs; the thesis with its shortcomings is mine.

Every bit of this work was made possible by the National Aeronautics and Space Administration on behalf of the people of the United States of America, through the NASA Graduate Student Researchers Program, under grant NGT 21-002-823. I do not know a nobler use for public money than supporting such an outstanding educational opportunity.

Without Dr. Michael J Mumma, Head of the Planetary Systems Branch of the Laboratory for Extraterrestrial Physics at the NASA Goddard Space Flight Center, discoverer of the natural planetary lasers at Mars and Venus, this project would not have occurred at all. The first to recognize and champion their potential, he performed the first performance assessments which justified this closer study. And in acting as the NASA Technical Advisor for this project, he sincerely and tirelessly answered my endless stream of naive "two minute" questions.

No one sees new views without standing on the shoulders of predecessors; the text documents my indebtedness to published authors in many fields. References to particular products are intended to recognize through use the often remarkable state of their art, not to endorse companies.

I thank Dr. Bruce K Donaldson, Professor of Aerospace Engineering at the University of Maryland and my Graduate Advisor, particularly for his courage and tenacity in sheltering unconventional work within relentlessly conventional surroundings, for preserving the freedom I needed to pursue that work independently, and for always keeping my professional development his primary goal.

I am most grateful to Dr. Michael A Coplan, of the Institute for Physical Science and Technology at the University of Maryland, for his enthusiasm in joining the Thesis Examining Committee, for his eager instruction and always sound advice, for starting me on the path to NASA funding, and for giving me a job upon my arrival at UMCP in 1984.

I thank Dr. Jean-Paul Richard, Professor of Physics and Astronomy at the University of Maryland, for his participation on the Examining Committee and in particular for his specific comments in that capacity.

I thank Dr. Bruce Berger, Professor of Mechanical Engineering at the University of Maryland, for his participation on the Examining Committee as well.

I thank Dr. Stuart Greenwood, of the University Research Foundation, for his unflagging interest in the project, his perseverance in reading the chapters as I produced them, and his perspective on extraterrestrial life.

I thank Peter Millard, Architect, for attending the Thesis defense, for his ruthless intelligence and lithic confidence in me, and for awakening in me at Yale the belief that the practice of design can be learned.

Without Joe Langdon, Reference Librarian of NASA Goddard's Homer E Newell Library, this work would be vastly less inclusive even than it is. Only with his expert help could I tap the enormous amount of stimulating information awaiting use.

Similarly, without Mike Green and particularly John Iarocci to help untangle operating intricacies of the UMCP Aerospace Series HP-1000 Model A-900 research computer, the graphed parametric studies would not be as they are.

Fellow NASA graduate students Tilak Hewagama and Jeff Goldstein, as well as other researchers at the Goddard LEP, helped repeatedly to clarify technical issues of real lasers.

I thank Michael D Griffin, the only true space professor in our department until he left, for his seminal instruction and encouragement.

I thank those who read or discussed portions of the work along the way, including Dr. Robert L Forward and Robert M Sherwood, for their perspectives.

B & B's special thanks go to Aaron A Salzberg, and grad student accomplices (the other) B & B, for tolerating constantly my need to be consumed by my work and talk about it incessantly, for appreciating directly the hazards of graduate life, and for all the friendship and extracurricular support which make the UMCP era memorable.

I owe my family immeasurably for supporting the idealistic goals they sent me to school to develop, as well as every tactical decision redefining those goals. Model rockets, fish breeding, microbiology, classical piano, art history, martial

arts --- what in the world phase are we in now? It's a long way from the Church Lane basement.

Last and first I am privileged to thank RW, without whose irresistible diversion, and unshakable confidence in this space cadet, the world would be bleak indeed. There's a universe out there waiting for us.

TABLE OF CONTENTS

Preface	ii
Nomenclature	xvii
Variables	xvii
Units	xxv
Prefixes, Subscripts, Superscripts, Symbols	xxvi
Acronyms	xxviii
Introduction	1
<u>PART 1: BACKGROUND</u>	6
Chapter 1 Defining the Mission	7
The Search for ExtraTerrestrial Intelligence	8
Communication with ExtraTerrestrial Intelligence	16
CETI Methods	18
References	24

Chapter 2	Controlling Lasers	28
	Lasers	29
	Laser Subtleties	31
	CO ₂ Lasers	35
	Planetary Laser Emission	39
	Controlling Light	42
	Nonlinear Optical Phase Conjugation	48
	Figures 2-1 and 2-2	55
	Appendix A2-1: Stimulated Brillouin Scattering	57
	A2-2: Four Wave Mixing	60
	References	65
Chapter 3	Optical Communication	72
	Digital Sampling	74
	Reconstitution and Errors	77
	Presumed Receiver	82
	Figures 3-1 and 3-2	89
	Appendix A3-1: Interstellar absorption	90
	A3-2: Doppler shift for interstellar lasers	91
	A3-3: The link equation	94
	References	97

PART 2 PLANETARY LASERS 99

Chapter 4 System Overview 100

The Orbiting Resonator	102
Extracting Output	109
Modulating and Transmitting	114
Assembly and Maintenance	120
Link Performance	124

Figures 4-1 through 4-29	126
--------------------------	-----

Appendix A4-1: Basic Vertex Station systems sketch	155
A4-2: Stations 2, 3, 4 and 5 performance	164
A4-3: 1δ & 1ϵ systems sketch	168
A4-4: 1δ & 1ϵ systems performance	177
A4-5: Transducer systems sketch	182
A4-6: Transducer systems performance	187
A4-7: Ring systems sketch	190
A4-8: Ring station-keeping performance	195
A4-9: Program LINKPERF.FTN	197

References	199
------------	-----

Chapter 5 Planetary Resonators 200

Ring Resonator Configuration	201
Sweeping Footprint	206
Doppler Shift	208
Ring Laser Theory	210

Figures 5-1 through 5-15	213
--------------------------	-----

Appendix	A5-1: Program MARS.FTN	224
	A5-2: Orthogonally sweeping gain	225
	A5-3: Intracavity Doppler shift	227
	A5-4: Ring laser lock-in threshold	228
References		229
Chapter 6	Planetology and Astrodynamics	231
	The Martian Laser	232
	The Venusian Laser	240
	Figures 6-1 through 6-10	247
Appendix	A6-1: Sun-synchronous orbits at Mars	256
	A6-2: Program MARSYEAR.FTN	258
	A6-3: Conventional planetary gravity model	263
	A6-4: Resonator altitude variations at Mars	270
	A6-5: Overall gain for resonators at Mars	272
	A6-6: Altitude variations at Venus	275
	A6-7: Astrodynamics for resonator at Venus	279
References		282
Chapter 7	The Optical Path	286
	Cavity Diameter	287
	Basic Vertex Stations	289
	Phase Conjugation	292

Mirrors	295
Frequency Selection and Output Coupling	302
Vertex Station 1	307
Loss Accounting	309
Cavity Control	315
The Output Beam	319
Libration Point Modulator Stations	324
Figures 7-1 through 7-14	332
Appendix A7-1: Proof of extreme ray congruence	346
A7-2: Micrometeoroid flux at Venus	348
A7-3: Mirror segment design	349
A7-4: The grating equation	352
A7-5: Proof of allowable grating orders	354
A7-6: CO ₂ rotational distribution	356
A7-7: Diffraction grating design	358
A7-8: Line tuning resolution	360
A7-9: Figure deviation for shallow curvature	361
A7-10: Pointing accuracy requirement	362
A7-11: Resonator pointing biases	363
A7-12: Oscillation linewidth	365
A7-13: Beam power	371
A7-14: Program APERTURE.FTN	373
A7-15: Program PWRDEN.FTN	374
A7-16: Beam-pointing to modulators	375
A7-17: Program ANGERR.FTN	376
A7-18: Mirror deflection for beam spread	377
References	378

Chapter 8	Spacecraft Control	382
Intelligent Structures		383
Processing		387
Control Approach		389
Perturbations		401
Actuators		411
Sensors		419
Appendix A8-1: Orbital perturbations due to Venus		428
A8-2: Gravity gradient torque		430
A8-3: Third-body perturbations		436
A8-4: Non-Keplerian orbit penalty		439
A8-5: Drag perturbation		442
A8-6: Radiation pressure effects		443
A8-7: Orbit eclipse and terminator impulse		446
References		448
 Chapter 9	 Spacecraft Systems	 455
Structure		456
Power		468
Thermal Management		480
Attitude Control		488
Propulsion		496
Maintenance		499
Figures 9-1 through 9-11		503
Appendix A9-1: Power radiator performance		512
A9-2: Low-grade radiator performance		515
A9-3: Estimated temperature highs		516

A9-4: Maximum AMCD rim speed	518
References	519
<u>PART 3</u> <u>CONTEXT AND MEANING</u>	529
Chapter 10 Non-Planetary Lasers	530
The Need for an Alternative	531
High-Power Space Lasers	533
Non-Planetary Laser Systems	537
References	545
Chapter 11 Cost	550
Meaningful Cost	551
Infrastructure Cost	553
Shortcut Technologies	559
References	567

Chapter 12	Interstellar Transportation	569
	Augmented Transfer Rate	571
	Biostasis and Molecular Recording	573
	The Transporter	577
	Figure 12-1	583
	References	584
Epilogue		586

NOMENCLATURE

Variables

The numbers in parentheses refer to specific chapter usages.

A	area (6) amplitude parameter
[A]	coordinate transformation matrix
a	acceleration ellipse semimajor axis (3) encoding base (6) planetary radius (9) albedo
B	link channel capacity (7) molecular energy series expansion 1st coefficient
B	magnetic induction
B _r	receiver channel bandwidth
B _s	source emission linewidth
b	ellipse semiminor axis
C _D	drag coefficient
C _{nm}	cosine coefficients in U expansion
c or c ₀	vacuum speed of light
c	(5) right triangle "adjacent" side
D _A	aperture diameter
D _c	cavity diameter
D _r	receiver diameter
D _s	diffracted Airy spot diameter

d	distance (5) diameter (7) groove spacing
E	(2) total EM field (6) eccentric anomaly (7) Young's modulus of elasticity (8) orbital specific mechanical energy
E	EM radiation electric vector
E_c	conjugate EM field
E_l	local oscillator field
E_p	probe EM field
E_r	source EM field
E_1, E_2	pump EM fields
e	orbital eccentricity
F	scalar force (3) flux (9) view factor
F_g	gravitational force
F_{gg}	gravity gradient force
F_s	separation force
f_c	carrier frequency
f_m	modulation frequency
f_r	recording frequency
f_s	sampling frequency
Δf	bandwidth
G	universal gravitational constant
g, g_o, g_E	Earth-normal gravitational acceleration
$g(r)$	gravitational acceleration field
H	scalar angular momentum
H_c	control momentum

h	Planck constant (3) signal pulse amplitude (5) hypotenuse (6) altitude of maximum tangential gain (6) scalar orbital angular momentum
h_s	orbital altitude
\hbar	$h/2\pi$
I	current
$[I]$	inertia tensor
I_e	insolation as function of orbital eccentricity
I_o	insolation as function of obliquity
I_{sp}	specific impulse
I_{xx}, I_{yy}, I_{zz}	moments of inertia
I_{xy}, I_{yz}, I_{xz}	products of inertia
I_*	local solar intensity
i	orbital inclination
J	molecular rotational state
J_n	zonal coefficients of spherical gravity model
K	amplitude SNR (6) additive parameter
KE	kinetic energy
k	Boltzmann constant
\mathbf{k}	light propagation vector
L	length
L_g or L'	effective gain length
l	luminosity
M	total mass (6) mean anomaly (7) atomic mass (8) primary mass

m	mass integer parameter (3) number of bits in unit group (3) visual magnitude (4) mass/area (6) secondary mass (7) diffracted order (7) inverse Poisson ratio
m_i	spacecraft initial mass
m_p	propellant mass
\dot{m}	mass flow rate
N	number of communicative civilizations in galaxy (3) discrete encoding levels
n	integer parameter (eg years, current loops, &c) (7) index of refraction (7) longitudinal cavity mode number
\mathbf{n}	unit normal
n_J	population at level J
P	power (2) polarization
$P(\cdot)$	probability of (\cdot)
$P_{nm}(x)$	Legendre functions
P_r	specific power intercepted
P_s	(3) specific intensity of source spot (10) source power
PE	probability of error
p	paraboloid parameter
\dot{Q}	heat flux
\mathbf{q}	sound propagation vector
R	radius (6) planetary radius

$R(r)$	radial component of U
R_c	cavity radius
R_E	radius of equilibrium
R_e	effective radius
R_s	radius of diffracted Airy spot
r	radius
	orbital radius
	distance
	(5) reflectivity
r	radius vector
r_a	apoapsis radius
r_e	(3) extinction ratio
	(6) equatorial radius
r_p	polar radius
	periapsis radius
r_s	scattered fraction
r, θ, ϕ	orthogonal spherical polar coordinate directions
S	area
S_{nm}	sine coefficients in U expansion
s	circular arc
	right triangle "opposite" side
T	absolute temperature
	(4,8) thrust
	(4) tether tension
T	torque
T_e	emissive temperature
T_s	temperature of deep space
T_P	orbital time period
t	time
	(4,7) thickness
U	gravitational potential
U_{eff}	effective U

u	exhaust exit velocity
V	volume
v	velocity
v_{cs}	circular orbital speed
W	weight
w	actuator stroke
x	distance (7) center-to-edge mirror deflection
x,y,z	orthogonal Cartesian coordinate directions
y	distance (7) deflection
Z	TE figure of merit atomic number
α	angle (7) incidence angle (9) Seebeck coefficient (9) thermal absorptivity
α_a, α_c	coefficients of crystal thermal expansion
α_L	longitudinal CTE
α_s	spacecraft thermal absorptivity
β	volume photon emission rate angle (7) diffracted angle
γ	angle (7) total system losses

Δ_R	receiver degradation factor
δ	altitude difference
ϵ	thermal emissivity
ϵ_R	received optical energy
ϵ_S	spacecraft thermal emissivity
ϵ_T	transmitted optical energy
ζ	damping ratio (7) body shielding factor
η	efficiency
η_d	diffraction efficiency
η_s	transmitter efficiency
θ	angle blaze angle
θ_s	solar zenith angle
$\Theta(\theta)$	latitudinal component of U
κ, κ_T	thermal conductivity
λ	wavelength
λ_D	Doppler-shifted wavelength
μ	gravitation parameter (GM)
ν	frequency of EM radiation (6) true anomaly
ν_{osc}	center frequency
$\Delta\nu$	emission (oscillation) linewidth
$\Delta\nu_d$	Doppler-broadened linewidth
$\Delta\nu_{osc}$	quantum-limited spectral width
π	pi (3.14159...)

ρ	density
Σ	sum of
σ	Stefan-Boltzmann constant (9) electrical conductivity (9) working stress
τ	(2) response time (3) bit period (3) bit integration time (5) time (6) EM propagation time around planetary resonator (7) coherence time
τ_c	cavity photon decay time
U_p	proper motion uncertainty
$\Phi(\phi)$	longitudinal component of U
ϕ	(6) phase angle (8) angle separating radius vectors
χ	EM susceptibility
Ω	rotation rate (2) sound frequency (9) orbital rate
Ω_L	lock-in threshold
$\dot{\Omega}$	longitudinal rate of orbital ascending node
ω	light circular frequency (4) angular rate (5,6) angle (6) frequency
$\Delta\omega_c$	cavity mode width
0,1	logical variables (OFF, ON)

Units

A	amp
AU	astronomical unit (semimajor axis of Earth orbit)
b	bit
cm^{-1}	wavenumbers or inverse cm
d	day (Earth day unless noted)
dB	decibel
deg or $^{\circ}$	angular degree
erg	energy (10^{-7} J)
eV	electron volt
g or gm	gram
Hz	hertz (1/s)
hp	horsepower
hr	hour
J	joule (Nm)
K	degree Kelvin
kg	kilogram
ly	light year ($9.46(10^{15})$ m)
MT	metric tonne (1000 kg)
m	meter
mag	magnitude
mgal	cm/s^2
min	minute
N	Newton (kg m/s^2)
Pa	Pascal (N/m^2)
p	per
pc	parsec (3.26 ly)
ph	photon
rad	radian
rpm	revolutions per minute
s or sec	second
sr	steradian
T	Tesla

tonne	1 MT
torr	pressure, mm of Hg (133.3 Pa)
V	volt
W	watt (thermal or electric, Nm/s)
We	watt (electric)
yr	year (Earth year unless noted)
\$	US dollar
"	inch (2.54 cm)

Prefixes

p	pico- (10^{-12})
n	nano- (10^{-9})
μ	micro- (10^{-6})
m	milli- (10^{-3})
c	centi- (10^{-2})
k	kilo- (10^3)
M	mega- (10^6)
G	giga- (10^9)
T	tera- (10^{12})
E	exa- (10^{18})
Δ	delta (change in)
d	differential
∂	partial differential

Subscripts

max	maximum
o	rest, nominal or reference
	parallel
⊥	perpendicular
☉ or *	sun
⊕	Earth
♀	Venus
♂	Mars
γ	orbital Station 1γ

Superscripts

T	transpose
*	excited species

Symbols

·	first derivative with respect to time
∫ _s	surface integral
∫ _c	closed path integral
()	vector
[]	matrix
∞	infinity
boldface	vector quantity or important numerical result

Acronyms

AAS	American Astronautical Society
ACOSS	active COSS
AD	anno Domini
AFWAL	Air Force Wright Aeronautical Laboratories
AGARD	Advisory Group for Aerospace Research and Development
AI	artificial intelligence
AIAA	American Institute of Aeronautics & Astronautics
ALNICO	aluminum nickel cobalt
AMCD	annular momentum-control device
AMTEC	alkali metal TE conversion
APCOSS	active and passive COSS
ATF	Astrometric Telescope Facility
BMEW	ballistic missile early warning
CBN	cubic boron nitride
CC	carbonaceous chondritic
C ³ I	command, control, communication & intelligence
CETI	communication with ET
CG	center of gravity
CM or cm	center of mass
CMG	control moment gyroscope
COAT	coherent optical adaptive technique
COSS	control of space structures
CP	center of pressure
CSI	controls-structures interaction
CTE	coefficient of thermal expansion
CVD	chemical vapor deposition
CW	continuous wave

DC	direct current or constant
DFWM	degenerate FWM
DLC	diamond-like carbon
DoD or DOD	Department of Defense
DoE	Department of Energy
DOF	degree of freedom
EM	electromagnetic
EMT	EM translator
EOL	end of life
ESA	European Space Agency
ET	extraterrestrial intelligence
FEM	finite element method
FIR	far IR
FWHM	full-width half-maximum
FWM	four wave mixing
GR	general relativity
HIP	hot isostatic pressing
HOE	holographic optical element
IECEC	Intersociety Energy Conversion Engineering Conference
IEEE	Institute of Electrical & Electronics Engineers
IF	intermediate frequency
IFF	if and only if
IOC	integrated optical circuit
IPL	interference phase loop
IR	infrared
IRAS	IR Astronomy Satellite
ISEE	International Sun-Earth Explorer

JPL	Jet Propulsion Laboratory
LAGEOS	Laser Geodynamic Satellite
LASER	light amplification through stimulated emission of radiation
LDEF	Long Duration Exposure Facility
LDR	Large Deployable Reflector
	Liquid Droplet Radiator
LEO	low Earth orbit
LQG	linear quadratic Gaussian
LSI	large scale integration
LSS	large space structures
LTE	local thermodynamic equilibrium
L1,L2,L3,L4,L5	Lagrange libration points
MCP	micro-channel plate
MHD	magnetohydrodynamics
MIT	Massachusetts Institute of Technology
MMC	metal matrix composite
MODFET	modulation-doped field-effect transistor
MPD	magnetoplasdynamic
MPP	Massively Parallel Processor
NASA	National Aeronautics & Space Administration
NATO	North Atlantic Treaty Organization
NEF	noise-equivalent flux
NFWM	nondegenerate FWM
NOPC	nonlinear optical phase conjugation
PCM	phase conjugate mirror
	pulse code modulation
PCR	phase conjugate resonator
PDT	precision diamond turning
PEMLM	photo-emitter membrane light modulator

PVD	plasma vapor deposition
PVDF	polyvinylidene fluoride
PZ	piezoelectric
PZT	lead zirconium titanate
RBMR	rotating bubble membrane radiator
RF	radio frequency
RGA	rate gyroscope assembly
RLG	ring laser gyroscope
RMF	reactive metal foam
rms	root-mean-square
RTG	radio-isotope TE generator
RZ	return-to-zero
SASS	six-axis space sensor
SBS	stimulated Brillouin scattering
SDI	Strategic Defense Initiative
SDIO	SDI Office
SETI	search for ET
SFC	Space Flight Center
SNR	signal-to-noise ratio
SPIE	Society of Photo-Optical Instrumentation Engineers
SPS	solar power satellite
SQUID	superconducting quantum interference device
SRS	stimulated Raman scattering
SSI	Space Studies Institute
ST or HST	Hubble Space Telescope
STS	Space Transportation System
TAARA	three-axis angular rate & acceleration sensor
TE	thermo-electric
TEM	transverse EM mode
TFE	THI fuel element

THI	thermionic
TPV	thermophotovoltaic
TSS	Tethered Satellite System
UHF	ultra-high frequency
US	United States of America
UV	ultraviolet
VCSS	vibration COSS
VE	visco-elastic
VR	vibrational-rotational
VRM	Venus Radar Mapper
WAHOO	a poker game
YAG	yttrium aluminum garnet

INTRODUCTION

Interstellar communication is a subject fraught with opinion, at times even vitriolic prejudice. For millennia, people patently presumed the existence of some form of life beyond Earth. In this age, however, when evolving technology rapidly continues to circumscribe the mysterious, contemplating our human place in the universe is not simple.

This work attempts to define rationally one method we might use to discover, or to establish, our place in the universe. Arthur C Clarke wisely cautions that alien life must be utterly different from us. Lacking therefore any just reason for ascribing motives or means to such life, we must if interested venture forth ourselves, either to find it or to become it. Being able to transmit complex, meaningful signals across the gulfs between stars could only aid our search for extrasolar meaning.

This treatise takes the form of an engineering feasibility study, examining prospects for using energies found in our solar system to accomplish efficient interstellar communication. Because of the extremely advanced nature of the problem, no one can yet declare with finality that the necessary abilities are either "feasible" or "unfeasible". My goal has been instead to provide a framework for viewing the problem, which each reader can use to evaluate for himself the evidence.

By exposing and treating vulnerable points, I hope to vaccinate readers, sensitizing them to respond critically to later treatments of this same problem, and in general to other glibly proposed, elaborate space systems. This study may therefore contain explicitly, like oncogenes, the seeds of its own conclusions' demise. If however it stimulates careful thinking about both the difficulty and promise of useful interstellar communication, it will have served well.

Taking on a speculative space design project of immense scale invites pervasive liabilities and rewards. An unusually dominant fraction of the "unknown" precludes both comforting technical detail and familiar references. But a lack of analogous precedents also exercises true systems integration, by enforcing that all facets receive a fresh look. And extreme space mission requirements ensure inventiveness, since solving new problems cleverly depends on a thorough exposure to advanced concepts.

A system design is a dense web, extensively cross-linked and interdependent. Its logic resides in its self-consistency, not necessarily in its chronological genesis. But reading is necessarily a serial activity. Thus although all parts of the project evolved simultaneously, influencing each other in iterative ways too numerous and subtle to record, I have arranged the thesis as a cycle of chapters. After an exposition of the subject, we embark on an odyssey through many realms of advanced technology, returning finally to the original subject with, I hope, a more informed and mature insight. A specific reference design, presented before the detailed systems analysis as a fait accompli, recurs as a compass along the way.

Part 1 (BACKGROUND) sets the scene for designing interstellar communication lasers with a trio of tutorial chapters:

Chapter 1 (Defining the Mission), by surveying contemporary thought about searching for extraterrestrial intelligence, derives an impetus for developing the ability to transmit large amounts of data over interstellar distances.

Chapter 2 (Controlling Lasers) discusses salient aspects of lasers, focusing in on carbon dioxide lasers and especially the natural atmospheric lasers of our solar system, and then surveys modern phase control techniques.

Chapter 3 (Optical Communication) establishes the utility of light as a signal carrier, then specifies important interstellar link parameters and derives a fundamental equation to model link capacity.

Part 2 (PLANETARY LASERS), the core of the work, designs in some detail one laser transmitter system for the interstellar mission, based on the Venusian natural laser:

Chapter 4 (System Overview) tours the components and operation of the complete design, analyzing its performance and orienting the discussions of five succeeding chapters.

Chapter 5 (Planetary Resonators) investigates astronomical constraints on establishing resonators to

extract continuous, steady, useful laser beams from planetary atmospheres.

Chapter 6 (Planetology and Astrodynamics) selects the better of our two available candidate planets (Mars and Venus) by comparing their lasing environments, and specifies a resonator orbit at Venus.

Chapter 7 (The Optical Path) transforms the physics of making and controlling a laser into engineering performance specifications for all the geometrical surfaces touching the beam, from genesis to transmission.

Chapter 8 (Spacecraft Control) outlines the sensors, actuators, information processing and artificial intelligence necessary to operate the laser fleet as a single apparatus in a changing environment.

Chapter 9 (Spacecraft Systems) selects technologies for active structure, power generation, thermal management, overall attitude control, propulsion and maintenance.

Part 3 (CONTEXT AND MEANING) projects the role of interstellar communication lasers in an advanced space culture, and their seemingly inevitable uses.

Chapter 10 (Non-Planetary Lasers) examines the major differences between planetary lasers and alternative kinds which could perform the same interstellar mission.

Chapter 11 (Cost) defines through its infrastructure the type of civilization capable of building and using

interstellar lasers, reviewing the effects of anticipated but unpredictable technical progress.

Chapter 12 (Interstellar Transportation) explains how efficient informational links among star systems would assist a nanotechnological culture in expanding rapidly through the galaxy.

Readers mainly curious about prospects for extraterrestrial intelligence, and possible human futures in galactic history, may be stimulated by Chapters 1, 11 and 12, and the Epilogue.

Readers interested primarily in why interstellar lasers could be important, and in what they might be like, should concentrate on Chapters 1, 4, 10, 11 and 12, and the Epilogue.

Readers looking explicitly for technical spacecraft design will find it in Chapters 4, 5, 6, 7, 8 and 9 (all of Part 2), and particularly in the appendices of those chapters.

Comprehensive appreciation of the complex challenges posed by the problem of efficient interstellar communication, and the range of options to be winnowed in solving that problem, requires studying all of Chapters 1 through 12, and the Epilogue.

PART 1

BACKGROUND

It is difficult to say what is impossible,
for the dream of yesterday is the hope of
today and the reality of tomorrow.

-- Robert H Goddard

CHAPTER 1

DEFINING THE MISSION

Chapter Abstract - Predictive efforts to determine the prevalence of extraterrestrial civilizations are academic. If interstellar travel is feasible, an irrepressible settlement wave could sweep the galaxy in a cosmically short time, establishing hegemony through occupational priority for a replicating lineage of civilizations. Various methods are available to search for evidence of the progress of that wave. Should a stellar culture desire informational contact with other star systems, laser wavelengths are at least as appropriate as radio, particularly for transmitting precisely targeted, elaborate signals. An infrared laser system based on demonstrated renewable solar system resources, and capable of large data transfer rates over distances out to 25 pc, would be useful both for establishing links with alien cultures and for maintaining contact with distant human colonies, should we ourselves eventually initiate an interstellar settlement wave.

The Search for Extraterrestrial Intelligence

Planning the modern Search for ExtraTerrestrial Intelligence (SETI) has, since its start a quarter century ago and until recently, centered on the Drake equation, intended to predict the prevalence of advanced communicative civilizations ("like" ours) in the Milky Way.

Appearing in as many variations as there are authors who rely on it, the equation attempts to derive the number N of recognizable, communicative civilizations in our galaxy by multiplying together a series of astronomical, biological and social probabilities. Assuming Copernican homogeneity, the equation reasons that of the $\sim 10^{11}$ stars in our galaxy, a certain fraction will have been stable, single suns of about the same size, luminosity, spectral class and age as our own G2-V dwarf. Of those, a further fraction will have been surrounded by planetary systems, including bodies of similar mass, composition and distance from their primary as Earth. Some of these Earth-analogs will presumably have developed life, and in some cases the life will have evolved society, then intelligence, then technology and finally a capacity and maybe even desire to communicate with life around other stars. Of such civilizations, some will exist right now, and some of those will be attempting contact. If we search and are lucky, some of those signals, finally, may happen to reach us while we are looking.

For a variety of reasons, it turns out that the Drake equation, albeit stimulating, is not really useful. Since almost every term in the product involves an uncertainty of several orders of magnitude, the choice of "optimistic" or "pessimistic" combinations of values (any of which is

theoretically defensible and all of which have fallen in and out of fashion over three decades) results in values of N which vary from over 10^9 , to 100, to 1 per 40 galaxies, to vanishingly rare [Hart, 80]. When used predictively then, the Drake equation as well as its most oft-quoted conclusion of $N \approx 10^5 - 10^6$, is unreliably arbitrary.

The equation's motive -- determining N -- can be approached more aptly, and with greater deductive validity, by examining the observable effects of different values rather than by speculating on their causes. We can do this by outgrowing the fundamental presumption of the Drake equation: that life must originate independently in all appropriate stellar systems.

Tenable belief during the Drake equation's ascendancy held that interstellar travel was impossible because the tremendous distances would require too-immense energies. Several studies later, however, we can already reasonably project several types of starships based on current physics and engineering [Dyson, 82]. At the fast extreme, diaphanous sailing probes could be accelerated by lasers [Forward, 84] or microwaves beamed from solar orbit, to 0.2 c speeds on unmanned voyages. Such craft could reconnoiter nearby star systems within one human generation, sending back analyses of their discoveries. Then, for instance, vast and heavy worldships using nuclear-electric or nuclear-pulse propulsion might journey at 0.01 c to eligible nearby systems. Such vessels would carry sufficient energy and material resources to sustain a replicating population of at least 500 people [Jones, 85] for the many generations such interstellar colonization would take.

These examples set only a lower bound on feasibility. Allowing any technical extrapolation broadens the starship array considerably. One modest design by Dyson [79] would have

people travel slowly through interstellar space inside itinerant comets displaced from the Oort cloud, whose surfaces collected the energy of feeble starlight with space-tolerant plantlife. An even wilder scheme would have human crews, themselves genetically engineered for longevity and diversity [Bracewell, 82], piloting catalytic hydrogen fusion ramjets at relativistic speeds among widely-scattered stars [Martin & Bond, 80] to "green the galaxy" with self-replicating von Neumann machines and tailored biota [Dyson, 79].

While these latter, less easily defended ideas represent approaches rather than solutions to the problem of interstellar transportation, we certainly cannot rationally preclude the potential of new discoveries and technical progress in this embryonic field (Chapter 11) -- indeed, a xerox machine would have seemed like magic to Gutenberg. But even disallowing those schemes requiring technologies we have not yet developed ourselves, there is as Dyson says "no lack of propulsion systems available to any creatures which possess...a desire to travel around in the galaxy." Creatures like us might possess that desire because stellar systems represent gravitationally collected lodes of matter and energy in an otherwise empty universe, kernels of order in the vast void of space around which entirely new civilizations might grow. Creatures like us might journey to other stars, given enough time, simply because they beckon across the emptiness.

If a stellar civilization (such as ours might become within a few centuries) built a few slow (0.04c) starships which took 2 - 3 centuries to travel to nearby stars, and even if those colonies took 7 - 8 subsequent centuries before launching their own expeditions, the ensuing settlement wave would grow outward at about 1 ly/century, given an average 10 ly step between stars. The startling result is that since it is $\sim 10^5$ ly across, the entire galaxy would thereby become

colonized in fewer than ten million years [Papagiannis, 80]. This interval, whose scale is rather insensitive to most particular choices of the colonization parameters [Drake, 80], is only about 0.1 % of the age of the galaxy, and only an order of magnitude greater than Homo sapiens' short time on Earth so far; cosmically, geologically and even biologically it is an insignificant span of time.

But would a succession of civilizations really colonize an entire galaxy? Without invoking teleology, four observable features of the order we call life are [Papagiannis, 80]:

"Life tends to expand to occupy all available space.

Life adapts to the requirements of every available space.

Life evolves...higher levels of organization.

The higher the level of organization, the faster [such organization] increases."

Colonization is the fastest, most complete way for life to fill the galaxy, a strong argument for its inevitability. Once a settlement wave started it would be virtually impossible to repress -- the most precocious civilizations would determine its replication rate. Just as every continent and island on Earth was inhabited by Homo sapiens long before each place could independently have evolved people, so must the entire Milky Way become infused with technological intelligence long before each star system could evolve it independently. The galactic infestation, graphed against cosmic time, must occur as a step function. Such an eruption leaves us, here now, with three possibilities. First, no one has begun colonizing yet. Second, the 10 million year infestation is happening around us now. Third, the expansion of life throughout the galaxy is history.

If no cultures have yet spawned progeny throughout the galaxy, N will be vanishingly small; our searches and signals will yield no results. Current "absence of verifiable evidence" tempts increasingly many to conclude that this first case must be fact. Whether or not life exists elsewhere, since it does exist here and since we can already envision the means to populate the galaxy, maturing into a stellar culture would assure us the evolutionarily advantageous position of primacy. Certainly under these conditions the Milky Way would end up at least shared by human descendants.

If the replication wave is moving just now, N will take on some rapidly increasing value. Most authors disallow this case because its probability is only 0.001, given the galaxy's age. This probabilistic proscription fails slightly when filtered through realistic analysis. First, advanced familiar life probably has not had the entire galactic age to develop. Our solar system, for instance, is only half the galaxy's age. Furthermore, we have no evidence at all that some event 4 - 5 billion years ago, such as a radiation burst from the energetic galactic center, did not reset the evolutionary clocks of all planets to within 10^4 years of each other [Troitskii, 80]. In such a case, allowing a variation of a billion years for technology to evolve increases to 1 % the chance of our being in the age of colonization, a small but nonzero probability which could indeed represent reality (somebody wins the lottery). If the wave is taking 10 million years to cross the Milky Way, we still might have the opportunity to take part in it, depending on exactly where and when it began.

If the populating wave is already over, then virtually every useful stable star system will be someone's found home and N will be of order 10^8 in our galaxy. Since our solar system is eligible, the Fermi Paradox is a natural question:

Where are they? "Their" apparent non-existence has, like our empty-handed SETI search to date, led to the consensus that we must be either alone, or one of the first technological species in the galaxy. Most astronomers who for "scientific" or "statistical" or (most likely) emotional reasons find it inconceivable that Homo sapiens could be first, have tried to invalidate the colonization scenario on economic, behavioral or physical grounds. The foundation of all such arguments is treacherous because even though the necessary restrictions might apply to some cultures, they could not apply to all cultures all the time. As noted earlier, the expansion rate of the settlement wave is set by the quickest replicators in it. Since less diversity in space even than on Earth is implausible [Hart, 80], if N were just about anything greater than unity, it would soon become huge.

N could be huge; in fact our solar system could easily have been colonized. We commonly and arrogantly presume that we are so inherently fascinating that interstellar travellers would contact us. However, in a universe full of life, a planet overrun with animals which poisoned their biosphere, and spent their resources on enough weaponry to eradicate themselves 20 times over, might not appear so attractive. It might seem best to leave such creatures to themselves, either to grow up or to succumb to natural selection. And in a truly crowded galaxy, interstellar immigrants might avoid contacting the indigenous inhabitants of a deep planetary gravity well until they had securely established their own civilization where the radiant energy and material resources were both optimally available -- the asteroid belt.

Our entire solar system might have been missed; the mathematics of theoretical ecology predicts substantial colonization gaps in an otherwise saturated galaxy [Turner, 85]. Whereas we can predict a general expansion of life

throughout the stars once a replication wave starts, we cannot speculate productively on its fine details; radically different evolutionary paths preclude coincidence of either type or ability. Life may well come from environments so hostile to us that we would overlook it [Shapiro & Feinberg, 82]. We can be certain that even our own spawn, separated from us by light years, eons, and willful speciation, will be alien [Dyson, 79]. Clearly, enough alternatives exist that the Fermi Paradox "cannot be considered a paradox at all" [Kuiper, 80].

If indeed life has not yet swept through the galaxy, N could be small. Although "stars with both ages and heavy element abundances comparable with those of the solar system are quite common in the galaxy" [Trimble, 82], and although basic life appeared on Earth almost immediately after its surface cooled and meteoritic bombardment abated [Papagiannis, 85], nothing at all suggests that the stability over billions of years which our planet has enjoyed is commonplace. We fortuitously orbit the galactic center in an inter-arm region, a safe distance away from frequent supernovae [Papagiannis, 85]. Our radiation-protective geomagnetic field has endured (although it periodically reverses). Earth has balanced on the water-based climatic knife edge between Venusian overheating and Martian freezing. And cataclysms, whether geological burps or cosmic peltings, have occurred just often and severely enough to stimulate rapid speciation. Current understanding reveals that, while our incubation stability may not be a unique or exclusive miracle, our existence is a miracle nonetheless.

Recently, the several authors' incendiary debate over what could be the real answer has melted resignedly into the awareness that only extensive and inclusive empiricism can settle the question. As Kuiper says, "Our knowledge about the present absence of...a galactic civilization is only as good as

the extent of our searches for it." We have not yet completely explored the likely places even in our solar system; infrared (IR) [Papagiannis, 85] or in situ inspection of the asteroid belt might reveal surprising swarms of vessels practically invisible from Earth. Decades ago, Dyson [63] proposed that a Type 2 civilization (defined by Kardashev as one which controls the output of an entire star [Dyson, 79]) would disassemble the planets of its star system, constructing biosphere elements around the star to intercept virtually all of its energy. Such a Dyson sphere would inevitably reveal its presence by its waste heat IR emission. But if we assume Type 2 civilizations use their energy efficiently, rejecting only low-temperature waste [Rood, 87], then their signature becomes indistinguishable from the spectra of stars surrounded naturally by dust and debris. It has been said hyperbolically that to IRAS (InfraRed Astronomy Satellite), all stars look like Dyson spheres.

And in the next section we see that, despite our theories, we do not know which electromagnetic wavelengths communicative civilizations might use intentionally to signal each other, to welcome newcomers, or to announce their own presence. So far we have engaged only the barest of sporadic searches, mostly microwave at 21 cm wavelength, adding up to less than 120,000 hr [Papagiannis, 85]. NASA is only now undertaking a concerted SETI effort (alas, presuming only microwaves so far). While it would take a very long time to convince ourselves that no one else is out there, verifying a large N could easily occur within the next century, perhaps much sooner. And considering the growing pains Homo sapiens suffers, contrasted with the immense survival motivation that either $N = 1$ or $N = 10^8$ would provide, SETI has obvious short-term value beyond assuaging an ancient curiosity.

Communication with Extraterrestrial Intelligence

If our searches found another stellar culture, we would probably try to contact it. Optimally, in fact, we would try ourselves to contact other star systems even without knowing whether they harbored intelligent life, because such Communication with ExtraTerrestrial Intelligence (CETI) would increase the chance of SETI success. Whereas a culture announcing its existence might like to broadcast omnidirectionally, it could target individual stars with narrow beams much earlier and more economically. However, if its transmission parameter space (choice of directions and intervals) does not overlap a searcher's parameter space after the lightspeed propagation delay specified by their separation, then contact cannot occur. Therefore a searcher who announces his location and interest maximizes the chance that someone else will send in his direction.

Carrying out a CETI project, we would probably send a repeating decoding tutorial interspersed (time-division multiplexed) with surveys of our science and our art, in video, audio and symbolic form. Most technological knowledge we sent would probably be already known at least approximately by beings able to decipher our transmission (Chapter 3). Our local solar system "natural history", however, and in particular its biology would prove most interesting to distant aliens because of its uniqueness, as might our cultural creations. A more technically advanced civilization might in turn send us back useful or even critical information, such as future chemistry and physics. Thus, beyond the undeniable intrinsic stimulation of its cultural exchange, CETI might directly affect our own development.

Exposing every detail of our knowledge to the universe introduces a unique uneasiness to our planetary-centered minds. Although Papagiannis [83] argues persuasively that natural selection would favor benign, stable stellar civilizations since the majority of each would necessarily be limited to its finite home system, it is not at all clear that cultures would extend that self-tolerance to other, expansionary cultures. Niche competition will probably occur galactically. Still, evolutionary survival is best served by multiplying oneself, not by eradicating competitors directly, and the energetics of interstellar travel would make invasion or repression, as we understand and fear them, uselessly formidable undertakings; primacy is much cheaper than war. An already occupied neighborhood is the most likely "danger" we might encounter.

Besides, powerful UHF carriers and Ballistic Missile Early Warning (BMEW) radar beams have already signalled our existence and location to a distance of over 20 ly (increasing obviously at 1 ly/year). Since these beams contain enough information [Sullivan, 80] for a listener to deduce the Earth's orbital parameters and rotation rate, a map of transmitters and estimates of their physical size, we have already leaked enough to enable voyeuristic hypotheses about our biology and abilities. Fear of possessive extraterrestrials may be unjustified; it is certainly moot.

Whether or not a planetary civilization rationalizes the expense of unrequited CETI, it might develop a need for intra-species interstellar communication anyway. For instance, the logic leading to an eventually large N includes the probability of our contributing to it, or even causing it altogether. Indeed unless our neighboring star systems are already occupied, it would appear certain that we will expand to them within the next thousand years. Assisting that growth will be our own interstellar information network.

Although each leg of the expansion might take centuries, the message delay between adjacent stars is only a few years. The only practical, extensive cosmic connection among these star systems populated by extraterrestrial humans will be informational. Therefore a real need must develop sooner or later for targeted interstellar communication systems capable of large data transfer rates. Such systems are the subject of this study.

CETI Methods

For sending messages to human receivers or other previously located civilizations, the issues of target choice and transmission duty cycle would be trivially determined by known factors. In fact, for both of these cases we can easily imagine wanting a full duty cycle, so that dedicated facilities would most likely be built to accommodate them (Part 3). For scaling purposes in designing an interstellar network among propagating human settlements, we may assume a worst-case separation on the order of 10 pc (33 ly) for each leg of the link, as repeater stations throughout the network could relay signals farther on.

Sending messages to potentially alive sites requires more speculative scale decisions, however. First, in order to have target sites at all we must assume that life would concentrate close to stars, reasonable since usable material resources are more compactly available at such gravitational foci than in the interstellar void. Consistent with the plan of sending to the same stars we might expect to receive signals from, we can take NASA's contemporary SETI program as an

appropriate model. A search parameter space may be defined explicitly as the product of several critical quantities: the number of targets, the total bandwidth of electromagnetic frequencies monitored, and the typical duration of scrutiny. Increasing the size of any of these dimensions enlarges the parameter space, a good thing if we want a successful search. To maximize its use of limited resources, NASA cleverly divides its effort into two intersecting parameter spaces of differing emphases, thereby spanning more of the total space set than any single scheme of comparable cost could. The first technique is an all-sky survey at constant flux level, intended to find any "obvious", bright, perhaps distant, major beacons [Gulkis et al, 80].

The second NASA technique consists of a targeted search at much higher sensitivity, to scrutinize likely candidate stars in our galactic neighborhood. Paradigmatic for our communication design, this effort singles out the 773 stars of luminosity class V and spectral types F, G and K catalogued within 25 pc (82 ly) of our sun [Seeger & Wolfe, 85]. K-type stars have been modeled to have no "habitable" zone (that span of orbital radii which permits planetary liquid water over stable billions of years) and thus appear unlikely as civilization birthstars [Hart, 79]. However, being attractive choices as adopted stellar homes, they should be included in any search or communication scheme which posits colonization.

The broadcast ranges of our two scenarios -- human settlement (10 pc) and true CETI (25 pc) -- are roughly within a factor of 2; therefore the numbers of target stars contained by their respective volumes are similar within an order of magnitude. Choosing the larger values as a reasonable upper limit, we establish a communication scale goal of 25 pc radius, enclosing of order 800 target stars. Whereas this

may seem a minuscule number in the face of a vast universe, in operational fact it comprises quite a challenge. If 800 stars were addressed by one CETI transmitter (which must accumulate downtime at least by slewing to each new target), it could devote only a few hours per year to each of them. Such a duty ratio, on the order of 10^{-3} , already stretches implausibly thin the probability of its transmission being noticed. Dividing the broadcast through a simultaneously multi-directed transmitter, on the other hand, would improve the duty ratio at the price, for a given signal system, of less transmitted power to each target. 25 pc is therefore in any case an appropriate upper limit.

The bulk of SETI literature, and in fact NASA's funded project, presumes interstellar communication will occur using microwaves. Many decades' prior experience manipulating microwaves led naturally to Drake's inaugural microwave SETI experiment almost three decades ago, an expedient choice bolstered then and since by rationalizing argument. Several studies repeatedly pointed to the 1 - 10 GHz range as the best compromise between absorption by the interstellar medium and obscuration by natural noise [Morrison, 85], figuring that the frequency which would go the farthest while requiring the minimum broadcast power would be any civilization's logical choice for CETI.

Discovery of the 21 cm hydrogen spectral emission line (Morrison calls it the "most abundant photon in the universe") and nearby OH lines prompted the romantic and persistent first-generation notion of cosmic "watering-hole" frequencies around which hydrophilic carbon-based galactic life would flock to socialize. A recent derivative suggestion proposes an entire cosmic alphabet based on the hydrogen line spectrum [Hoang-Binh, 85]. In past analyses, vested exuberance induced some experts to claim, for instance, that "no laser system can

ever hope to compete with microwave systems...[if] microwaves had only recently been discovered, [they] would be hailed as the long-sought answer to interstellar communication" [Oliver, 74].

Like the Drake equation, however, the foundation of microwave dominance for SETI is in the end shaky. As Betz [87] points out, merely minimizing noise is a spurious goal, since improving communication efficiency depends really on maximizing the received signal-to-noise ratio (SNR). Assumptions about noise sources depend critically on whether a broadcast or targeted system is considered, and the SNR depends further on what kind of detection scheme is used. The much vaunted 21 cm "line", for instance, is both "poorly defined in frequency position" and "a place of high noise power" [Morrison, 85]. Perhaps the most common photon in the universe is too common. The essential reasons for preferring microwaves (maximizing use of familiar technology and minimizing required broadcast power) fade to anthropocentric artifacts in the light of a rigorous parametric comparison [Townes, 83]. If, as Townes says, "we have no assurance the microwave region is the one of choice for a civilization trying to communicate", then our search effort is best spent in diverse approaches.

By extension, we should develop equally diverse methods for CETI, particularly since by analyzing them carefully we can gain a much more mature perspective on promising SETI approaches. Some of the conclusions of this study, in fact, may explain the apparent cosmic silence; so far we have not even planned the type of search (IR), in the proper location (small solar orbits), required to detect a sophisticated (highly modulated) incoming CETI signal. According to Townes, "infrared is as good as, and may be a more favorable region...than the microwave region on the basis of reasonable assumptions." No region is favorable without the means to

realize its use, however. Consequently the purpose of this project is to establish, by designing a transmitter system, the feasibility and utility of using infrared lasers as tools for substantive interstellar data transfer.

Infrared lasers are an attractive CETI alternative for several practical reasons. Two advantages of laser light for interstellar communication, which subsequent chapters develop in detail, are that it can carry more information per unit time than microwaves (its frequency and hence available modulation bandwidth is much greater), and its highly directed nature fits aptly the problem of linking point targets informationally across space. IR wavelengths require less stringent optical precision (by a factor of roughly 15 - 25) than the visible laser wavelengths, and so represent the most practical to work with in the optical frequency region of the electromagnetic spectrum. Thus an interstellar IR laser would be certain both to advance the state of the art and to benefit from the large bandwidth available to optical carriers.

In addition, many natural galactic sources of IR radiation exist, with the direct consequence that those wavelengths would probably be often and thoroughly studied by a spacefaring culture. Such scientific monitoring would improve the chances of a CETI signal being noticed, particularly since the observed spatial coupling of an extraordinarily bright, spectrally narrow IR source with an otherwise optically boring star would appear peculiar. Finding the signal to be modulated artificially would then confirm its intent and origin.

Although a variety of candidates exists for even infrared space-based laser systems (Chapter 10), this study concentrates first on developing one of the most startling renewable natural resources proffered by our inner solar system: solar-pumped CO₂ laser emission at 10.6 μ m wavelength in the mesospheres of

Mars and Venus. Mumma [83] has proposed configuring a space-based resonator apparatus, consisting of essentially mirror spacecraft, to tap and use such enormous planetary gain media for the purpose of CETI. Certainly operating a solar-pumped planetary laser to probe the galaxy informationally previews a future maturity, by embodying embryonically both the technical skill and conceptual elegance required to evolve from a planetary into a stellar civilization. In this work, we develop the requirements, specifications, performance, and implications of a planetary laser system, accepting as its mission Communication with ExtraTerrestrial Intelligence, whether that extrasolar intelligence be alien and unknown, or our expatriated own.

References

C W Allen- Astrophysical Quantities 3rd ed (Athlone, 73).

Al Betz- "IR Laser Search" delivered at: NASA Space Life Sciences Symposium Washington DC (26 Jun 1987).

Roy L Bishop (ed)- Observer's Handbook 1986 (Royal Astronomical Society of Canada, 1986) data from: [D Hoffleit & C Jaschek- The Bright Star Catalogue (Yale University Observatory, 1982)].

R N Bracewell- "Preemption of the Galaxy by the First Advanced Civilization" in: [Hart & Zuckerman, 82].

A G W Cameron (ed)- Interstellar Communication (Benjamin, 1963).

Arthur C Clarke- "Visions of Space" Spaceflight Vol 28 (May 1986).

Frank Drake- "N is Neither Very Small Nor Very Large" in: [Papagiannis, 80].

Freeman J Dyson- "Search for Artificial Stellar Sources of Infrared Radiation" in: [Cameron, 63].

- Disturbing the Universe (Harper & Row, 1979).

- "Interstellar Propulsion Systems" in: [Hart & Zuckerman, 82].

- Robert L Forward- "Roundtrip Interstellar Travel Using
Laser-Pushed Lightsails" J Spacecraft and Rockets
Vol 21:2 (Mar-Apr 1984).
- Samuel Gulkis, Edward T Olsen, Jill Tarter- "A Bimodal Search
Strategy for SETI" in: [Papagiannis, 80].
- Michael H Hart- "Habitable Zones about Main Sequence Stars"
Icarus Vol 37 p 351-357 (1979).
- "N is Very Small" in: [Papagiannis, 80].
- Michael H Hart & Ben Zuckerman- (eds) Extraterrestrials:
Where Are They? (Pergamon 1982).
- D Hoang-Binh- "On a Cosmic Alphabet" in: [Papagiannis, 85].
- Sebastian von Hoerner- "Manifestations of Advanced Cosmic
Civilizations" in: [Papagiannis, 80].
- Eric M Jones- "Where Are They? Implications of Ancient and
Future Migrations" in: [Papagiannis, 85].
- S A Kaplan- [op cit Chapter 3].
- T B H Kuiper- "Galactic-Scale Civilization"
in: [Papagiannis, 80].
- A R Martin & A Bond- "Starships and Their Detectability"
in: [Papagiannis, 80].
- Philip Morrison- "Entropy, Life, and Communication"
in: [Ponnamperuma & Cameron, 74].

- "The Number N of Advanced Civilizations in Our Galaxy and the Question of Galactic Colonization" in: [Papagiannis, 80].

- "Twenty-Five Years of the Search for Extraterrestrial Communications" in: [Papagiannis, 85].

Michael J Mumma- [83] [op cit Chapter 2].

Bernard M Oliver- "Technical Considerations in Interstellar Communication" in: [Ponnamperuma & Cameron, 74].

Michael D Papagiannis (ed)- Strategies for the Search for Life in the Universe (Reidel, 1980).

- "The Number N of Galactic Civilizations Must Be Either Very Large or Very Small" in: [Papagiannis, 80].

- "Colonies in the Asteroid Belt, or a Missing Term in the Drake Equation" in: [Hart & Zuckerman, 82].

- "Natural Selection of Stellar Civilizations by the Limits of Growth" IAF Congress, Budapest Hungary (10-15 Oct, 1983).

- (ed) The Search for Extraterrestrial Life: Recent Developments International Astronomical Union, Proc. 112th Symp. Boston, June 1984 (Reidel, 1985).

Cyril Ponnamperuma & A G W Cameron- Interstellar Communication - Scientific Perspectives (Houghton Mifflin, 1974).

Robert T Rood- "Searching for Unicorns and Extraterrestrial Civilizations" Scientific Colloquium NASA Goddard Space Flight Center (15 May 1987).

Carl Sagan (ed)- Communication with Extraterrestrial Intelligence, Proc. Conf. Yerevan USSR, 1971 (MIT, 1973).

Charles L Seeger & John H Wolfe- "SETI: The Microwave Search Problem and the Targeted Search Approach" in: [Papagiannis, 85].

Robert Shapiro & Gerald Feinberg- "Possible Forms of Life in Environments Very Different from the Earth" in: [Hart & Zuckerman, 82].

W T Sullivan III- "Radio Leakage and Eavesdropping" in: [Papagiannis, 80].

C H Townes- "At What Wavelengths Should We Search for Signals from Extraterrestrial Intelligence?" Proc. Natl. Acad. Sci. USA Vol 80 p 1147-51 (Feb 1983).

Virginia Trimble- "Nucleosynthesis and Galactic Evolution: Implications for the Origin of Life" in: [Hart & Zuckerman, 82].

V S Troitskii- "A New Approach to the Number N of Advanced Civilizations in the Galaxy" in: [Papagiannis, 80].

Edwin L Turner- "Galactic Colonization and Competition in a Young Galactic Disk" in: [Papagiannis, 85].

CHAPTER 2

CONTROLLING LASERS

Chapter Abstract - Laser oscillation produces highly coherent and monochromatic light, useful for long-distance communication. Natural CO₂ laser emission has been observed and modeled in the mesospheres of Mars and Venus, suggesting the possibility of engineering space systems to take advantage of such large gain media for operating interstellar transmitters. Setting up the conditions necessary for lasing, and for employing an output beam effectively, include carefully controlling the phase properties of the circulating field. Active techniques for phase control involve complex closed-loop hardware. Passive, nonlinear optical phase conjugation techniques have been demonstrated both capable of enabling laser operation in otherwise unsatisfactory situations, and applicable to CO₂ lasers.

Lasers

The laser (Light Amplification through Stimulated Emission of Radiation) depends on physical principles worked out classically by Einstein as early as 1917. Rather than inhabiting a smoothly continuous energy spectrum, atoms and molecules occupy discrete (quantized) energy levels called eigenstates, whose values can be predicted using quantum mechanical theory. Such systems can absorb or emit only quantized amounts of energy equal to the discrete energy differences between eigenstates. Consequently the energy given off by an atom in "falling" from a higher to a lower state is precisely the amount it must absorb for the reverse "upward" transition.

Such energy commerce occurs by both dynamical and radiative mechanisms. Molecules of a gas, for instance, can transfer energy among themselves by colliding and glancing off each other's electrostatic fields. Alternatively, they can absorb or emit electromagnetic energy (photons) spontaneously, whose radiation frequency is related directly to the system transitional energy through Planck's constant. Because any given molecule or atom exhibits a unique set of possible transition signatures, this radiative transfer underlies such diverse and important phenomena as spectroscopy, the photoelectric effect, and lasers.

Siegman [86] explains that it is more accurate in terms of quantum theory to discuss populations of atoms or molecules than to attempt to picture them individually. The "instantaneous quantum state of any one individual atom is usually a time-varying mixture of quantum states", such that atoms predominantly occupying a lower energy level will "evolve" toward a higher level when supplied with energy from an external source. If a population is bathed in a field of

radiation whose frequency corresponds to a relevant transition energy, then the individuals in that population will be induced by the energy field to make that transition, both up and down, in phase with the stimulating photons. Furthermore, the probabilities, and therefore the rates, of this stimulated absorption and stimulated emission are identical, and proportional to the local field intensity.

Lasers exploit directly the properties of quantum transitions. If the normal energy distribution of a population, dominated by the lower eigenstate, is artificially inverted by some energy pumping mechanism, the resulting top-heavy population inversion acts as a quantum energy reservoir. Spontaneous emission will then stimulate further emission which, fed back, will grow exponentially until the inverted population is depleted---that is, until the energy distribution equilibrates such that stimulated emission equals stimulated absorption, which is called gain saturation. The cascade of emitted photons must be both in phase with the radiation field and monochromatic (to first order), since only one radiation frequency derives from the transition involved. As long as the external pump continues to populate the upper eigenstate, though, the gain medium will lase.

Consider an active laser gain medium placed within an optical resonator cavity, a device comprised most simply of two mutually facing mirrors (a Fabry-Perot etalon). As the circulating field oscillates in this cavity, it intensifies because the photons whose emission it stimulates add to it, monochromatically and in phase. If some transmissivity is purposely allowed one of the resonator mirrors, a small portion of the circulating field will leak out all the time. This emergent laser beam which has been coupled out of the resonator is typically highly coherent (because of its stimulated origin), monochromatic (because of the narrow frequency

lineshape resulting from the quantum transition) and therefore close to diffraction limited (it self-interferes only because it emerges through an aperture of finite dimensions). Seen as an energy-conversion device, then, the laser tames and organizes incoherent energy into a useful kind of light which can measure, illuminate, push, heat, cut, and vaporize materials, or carry messages. Its highly directed and spectrally pure nature makes the laser especially useful for transmitting dense messages over large distances.

Laser Subtleties

The quality of laser light, as just noted, is established by its coherence and its monochromaticity. The extreme size of a planetary laser enforces a clear connection between these two properties (Chapter 7), but for a simple review of most lasers they can be treated separately. Spatial coherence means that the circulating field's phase is constant across any plane section normal to the beam axis. This is a measure of the beam's transverse power-distribution homogeneity. Temporal coherence means that the field's phase at any given station along its axis remains constant. Thus the circulating field is typically a standing wave, resonant with the exact cavity length. Another way of saying this is that the laser radiation field is in phase with itself; it must repeat its electromagnetic structure exactly with each reflection for stimulated emission to be coherent.

Monochromaticity is the measure of how specific the laser's frequency (color) is. The frequency associated with any given transition is not ultimately precise, but rather

occupies a narrow but nonzero frequency band. The lineshape function which describes this frequency distribution is verifiably the same for upward and downward transitions, however. Two distinct classes of phenomena cause the lineshape function to spread. Homogeneous broadening results from features common to all the atoms or molecules in a medium, such as the nonzero radiative interaction time itself, and phase interruptions caused by acoustic energy (in solid crystals) or collisions (in high-pressure gases). Inhomogeneous broadening occurs because the transition frequency of each atom or molecule in the system is unique, due to local crystal irregularities (in solids) or molecular motion (in gases). A relevant example is the molecules of a gas laser, whose individual emitted frequencies get Doppler-shifted by an amount proportional to the molecule's axial speed within the cavity (nonzero in general for any gaseous system).

A broadened lineshape function allows oscillation at several evenly-spaced frequencies, the cavity modes, dependent on the resonator length. The total optical field strength of such a laser is the summation of all these modes, and therefore fluctuates due to their phase interferences, degrading the temporal coherence of the outcoupled beam. Optimally we would want only one mode to oscillate, but a practical solution for conventional lasers is mode-locking, in which the intensity fluctuations are constrained to be regular. One method of accomplishing this is to include in the medium a saturable absorber. Such a material becomes more transparent at higher optical intensities, thus favoring the mode with the highest gain.

A variety of pumping methods can maintain the quantum population inversion necessary for lasing. The simplest conceptually is direct, coherent pumping by another laser, in which the pumping photons match by virtue of their frequency

some transition gap of a secondary lasant. Direct incoherent pumping works the same way except that the pumping source, which is not a laser, provides a mixed-phase optical field of the proper frequency. Indirect incoherent pumping is another optical method in which the lasant becomes non-selectively excited, but after cascading down their "energy ladder" its atoms or molecules tend to collect in their long-lived upper laser level, from which they can then be stimulated to emit. The earliest solid-state ruby lasers, excited by xenon flashlamps, worked in this way. Another kind of indirect, incoherent optical pumping uses concentrated light to heat an intermediate blackbody, whose thermal emission spectrum peaks at the desired transition wavelength and bathes the lasant.

A common non-optical method for pumping the population inversion uses an electric discharge, whose electrons serve as energy carriers by colliding either with gaseous lasant molecules or intermediate metastable energy-storage molecules which in turn excite the lasant. Electrons are also the pumping agents in solid state semiconductor lasers. Another collisional method, employed by gasdynamic lasers, is to expand the lasant gas rapidly through a supersonic nozzle, ultimately exchanging system kinetic energy for laser upper-state excitation. Finally, chemical reactions can produce active excited species, and if replenished can result in continuous lasing.

Many loss mechanisms affect the operation of any real laser. To begin with, no resonator mirror is perfectly reflecting, so it must dissipate energy by absorption and scattering. Then too, the laser medium itself absorbs and scatters a small portion of the circulating field, both because the medium is impure and inhomogeneous at some scale, and because other undesired transitions coincidentally occur. And since a beam reflected from a resonator mirror of finite size

will spread slightly, diffraction losses occur on every pass as some of the circulating field "spills over the edges" at each next reflection. Finally, removing power from the circulating field when coupling out a useful beam constitutes a major loss. Obviously for laser oscillation to commence, single-pass gain must exceed single-pass total distributed losses. As the circulating field intensifies within the resonator cavity, gain saturation increasingly limits further amplification until the saturated gain balances system losses (including coupling loss), resulting in steady-state laser oscillation.

Both diffraction loss and coupling loss can be used to advantage. Inserting an intracavity aperture smaller than the cavity diameter, for example, reduces the amount of energy distributed into many different transverse resonator modes, thus improving spatial coherence. Also, by reducing the effective output coupler area, an internal aperture will increase the ratio of circulating field power to output power and therefore limit the portion of distributed losses budgeted to coupling. Both effects can also be realized through the use of an unstable resonator, in which both mirrors are convex toward each other. Such a configuration has only one ray trace that will not eventually "walk out" of the cavity upon successive round trips, resulting in a high-quality, exceptionally narrow beam.

A technique to increase peak power intermittently spoils, either mechanically or electro-optically, the reflectivity of one mirror, thus permitting the stored field strength to increase well beyond its normal saturated-gain value. Restoring reflectivity induces rapid oscillation and the sudden release of a high-power burst of laser light. Called Q-switching, this produces pulsed powers exceeding by orders of magnitude the normal continuous-wave (CW) operating power. High CW powers can be achieved by constantly replenishing the

depleted gain medium with fresh, excited material, as is commonly done in gasdynamic lasers and chemically pumped lasers.

CO₂ Lasers

Whereas the energy state of an atom is specified by its electronic structure, a multi-atomic molecule has in addition rotational and vibrational energies. All three types are quantized, taking on only discrete eigenvalues; transitions within these sets of energy levels correspond to characteristic emission and absorption frequencies in distinct spectral bands. Electronic transitions range in energy from 1 to 10 eV; pure rotational transitions extend from microwave ($\lambda \sim \text{cm}$) down to mid-infrared ($15 \mu\text{m}$) wavelengths; vibrational-rotational or VR transitions are at infrared (~ 30 to $2 \mu\text{m}$) wavelengths. Because of this spectral segregation, a simple model (the Born-Oppenheimer approximation) assumes that the quantum-mechanical wavefunction can be factored into the product of decoupled rotational, vibrational and electronic wave functions. The three types of energy are taken as independent, their effects merely additive.

The relevant quantum energy structure for CO₂ consists of gross vibrational levels overlaid by finely-spaced rotational levels (Figure 2-1). A nominally collinear triatomic molecule, CO₂ exhibits three basic, or normal vibration modes: symmetric stretching, bending, and asymmetric stretching. To first order the three are independent, so the molecule's vibrational state is conventionally represented by a triplet listing the respective quantized excitation of those three modes. Thus

(000) represents the ground state. Superimposed on this structure is the molecular rotational state, represented by a positive integer J . The highest-energy normal mode, (001), allows only odd values of J , whereas the intermediate levels (100) and (020) allow only even values of J .

These latter levels happen to be so extremely close in energy that they space themselves farther apart than expected. The wavefunction thus consists of contributions from both; because of this mixed state of so-called Fermi resonance, they are quantum-mechanically indistinguishable and usually designated together as (100,020)_{I,II}. Many subtly different transitions are possible between vibrational states. Transitions for which J remains the same (excluded for these CO₂ levels because of the odd/even discrepancy) belong to the Q-branch. Transitions for which the (001) (upper, odd) J is one greater than the lower belong to the R-branch, and those for which the lower J is one greater belong to the P-branch.

CO₂ is the archetypal gas laser medium, the 3-level behavior of which has been extensively studied, modelled and used practically. Absorbing a photon with wavelength near 4.23 μm enables a CO₂ molecule in the ground state (000) to jump to (001), as diagrammed in Figure 2-1. Subsequent transitions back down to intermediate levels emit photons with less energy; dropping to (100) yields a wavelength near 10.4 μm , depending on the exact J , while dropping to (020) yields a wavelength near 9.4 μm . In a simple laser cavity, the P-branch of the 10.4 μm band will compete successfully for the downward transitions at the expense of the R-branch and both branches of the 9.4 μm band. Emission will be gain-narrowed within the band to that J line with the lowest threshold, determined by the gas temperature (see Appendix A7-6 and Figure 7-4 for a specific example). Given all the possible

lasing frequencies, it is customary to speak of CO₂ lasing at 10.6 μm .

Because the laser levels for CO₂ are so close to the ground state, all its transitions are of the same order of magnitude. Thus the ratio of emitted photon energy to input excitation energy is a substantial fraction of unity and CO₂ has a high quantum efficiency. (This can be seen simply by noting that both the pumping and laser photons have comparable (infrared) wavelengths.) Additional factors, such as the relative number of molecules which actually make the downward laser transition, and how good the pumping mechanism is at getting them into the excited state in the first place, determine the laser's overall or plug efficiency. This practical limit varies from about 1 % for conventional solid-state lasers to near 100 % for GaAs junction lasers.

For several reasons, the CO₂ laser scores a hefty 30 % plug efficiency. Because the medium is optically thick (absorbing) at the 4.3 μm wavelength, photons emitted by radiative (non-lasing) decay of the upper (001) state are efficiently reabsorbed, keeping that state populated. The transition probabilities for 10.4 and 9.4 μm emission, normally three orders of magnitude less than that for the radiative 4.3 μm photons, thus dominate. Operationally then, the upper state is extremely long-lived (about 3 sec), so that molecules excited to a great variety of energy states cascade down into it and stay there until stimulated to emit.

Other reasons for the high plug efficiency of CO₂ lasers derive from exploiting fortuitous energy relationships which CO₂ shares with other admixed gases. The vibrational eigenstates of all homonuclear diatomic molecules are metastable, and can therefore act efficiently as tools to excite other molecules through collision. The energy gaps

between most of the first seven vibration levels of N_2 , and in particular the lowest, match almost exactly the (000) to (001) CO_2 transition (see again Figure 2-1), meaning that with a slight decrease in kinetic energy to make up the difference, N_2 can relax by pumping CO_2 into its upper laser level. Because at the typical operating pressure of a CO_2 laser (a few torr) most N_2 molecules lose their first-vibrational-state energy by colliding with CO_2 , N_2 is practically always used to pump such a laser.

Small amounts of He and H_2O help return the CO_2 from its lower laser level (after stimulated emission) back to the ground state quickly, thus reducing the turn-around time for each molecule and contributing greatly to system efficiency. He atoms also encourage oscillation on one CO_2 rotational line, by rethermalizing (filling in collisionally) the defect in the (001) level's Boltzmann distribution left by lasing. In lasers pumped by electron discharge, He atoms also transport excess heat to the cavity walls for conductive removal, and moderate the electron temperature of the discharge itself. Several different (and normally proprietary) gas mixtures are commonly used in such CO_2 lasers; one reference example is 23.5 % CO_2 , 12.7 % N_2 , 57 % He, 6.3 % Xe, and 0.5 % H_2 , at a total pressure of 16 torr. Some of the most interesting low-pressure CO_2 lasers, however, are neither experimental nor commercial (yet).

Planetary Laser Emission

Natural astronomical masers (microwave amplifiers) are observable in many interstellar clouds and circumstellar shells, where as long as $h\nu \sim kT$, population inversions pumped by photons, collisions and chemical recombination can occur relatively easily among closely-spaced rotational quantum levels. The higher-energy vibrational and electronic inversions needed for near-optical and optical lasing respectively would seem in general rarer, perhaps requiring a more fastidious pumping environment. Indeed, the first known natural lasers have only recently been discovered, operating in the atmospheres of Venus and Mars [Mumma et al, 81]. They have been extensively studied [Deming et al, 83] and modeled [Deming & Mumma, 83] since then, and independently confirmed [Gordiyets & Panchenko, 82] and discussed [Stepanova & Shved, 85].

Figure 2-2 compares graphically the primary atmospheric constituents of the major terrestrial planets. Not surprisingly, CO₂ dominates the mixtures on the two planets hosting natural lasers. Venus' is 96 % CO₂ and 3.5 % N₂ (by molar fraction), with gaseous traces of H₂O, SO₂, Ar, CO, Ne, HCl and HF, at a total surface pressure 90 times greater than Earth's. The troposphere (lower layer) includes a permanent planetary cloud layer of H₂SO₄ droplets between 50 - 80 km altitude, near the top of which blow 100 m/s winds. Mars' atmosphere is 95 % CO₂, 2.7 % N₂ and 1.6 % Ar, with gaseous traces of O₂, CO, H₂O, Ne, Kr, Xe and O₃, at a total surface pressure only 0.007 of Earth's. Most of Mars' H₂O is locked up as ice in polar caps (under seasonal CO₂ caps) and subsurface permafrost in regolith, perhaps even to fairly low latitudes. Its troposphere features thermal-tidal winds which periodically grow into global storms due to thermal feedback provided by the surface dust they suspend, resulting in a 50 km thick planetary blanket lasting for months. H₂O ice

clouds typically range to 25 km at the winter pole, and CO₂ ice clouds form as high as 60 km anywhere over the planet [Beatty et al, 82].

Above these contrasting regions of turbulent weather lie the calmer, rarefied and similar mesospheres of both planets, where their natural lasers operate. We focus the discussion for now on the better-documented Martian laser. Using an Earth-based infrared heterodyne spectrometer capable of resolving spatially a dimension roughly 1/8 the angular size of Mars, the discoverers of the planetary laser had measured intensity profiles to 5 MHz accuracy of several ro-vibrational lines in the 10.4 and 9.4 μm bands of CO₂, in order to study their known strong nonthermal emission spectrum. What they found on the 10.33 μm (967.7072 cm^{-1}) R8 line, for instance, was radiated energy $\sim 10^9$ times larger than the R8 would show if the Martian mesosphere obeyed Local Thermodynamic Equilibrium (LTE). That is, a >50:1 population inversion, with consequent gain amplification, exists in the lower Martian mesosphere; radiative relaxation reduces it to $\sim 7:1$ in the upper mesosphere. The emission peaks at an altitude of about 75 km (130 km at Venus), is proportional to incident solar flux, and vanishes on the dark side. Clearly the sun continually pumps CO₂ lasers at these planets.

The observed natural flux seems to result from the 10 μm radiative decay of CO₂ after its excitation by two processes [Deming et al, 83]. The first and major route is molecular absorption of near-infrared solar photons at many frequencies, followed by collisional transfer to the long-lived (001) state. The other significant route is direct pumping of the (001) state by 4.23 μm solar photons, a process which dominates at higher altitudes. The atmosphere there is optically thick enough at that wavelength that the molecular

absorption profile cannot deepen any more, but must broaden instead; thus the eligible absorptive molecules include those which are more and more Doppler-shifted by their kinetic velocity away from the rest absorption frequency ν_0 . These "fast" excited CO₂ molecules collide, redistributing their energy and slowing so that they re-emit closer to ν_0 . But because the medium is optically thick, these photons are quickly reabsorbed by other molecules and are thus said to be radiatively trapped. Since collisional relaxation to the unexcited state is inefficient at this altitude, the population remains inverted.

Model atmosphere temperature profiles were chosen to bracket Viking lander data, and theoretical thermo-quantum behavior of the gaseous composition calculated; although the total emergent flux depends strongly on temperature, the altitude of peak emission does not. The 120 K Mars model shows a maximum emitted intensity at large zenith angles (meaning long, tangential lines-of-sight through the mesosphere) over 20 times greater than that emitted in the zenith direction, and an almost linear dependence on the cosine of the angle between solar incidence and the zenith. The emission peaks when the line-of-sight minimum altitude is 66 km (130 km at Venus), at which point the optical depth (a normalized measure of the nominal travel length before a photon gets absorbed) is -0.07 for both planets, indicating gain. That means that one photon in 14 traveling this long path will produce another photon by stimulating emission, and that 4 % of the emergent photons are produced in this way. The models of Gordiyets and Panchenko [82] independently confirm these findings; they calculate a "radiation intensification coefficient" of from 1 - 4 (10⁻⁹) cm⁻¹, resulting in a single-pass gain of up to 10 % for the tangential subsolar path at 75 km Martian altitude, and between 2 - 40 % for a similar path at 130 km Venusian altitude.

Deming & Mumma's model is generous both by including almost every possible pumping line and in making some other assumptions, but the actual observed Venusian flux is still 74 % of the theoretical value. The Martian flux is 100 % of the theoretical value, hinting that some other processes, such as collisional deexcitation with H₂O vapor, may be helping there. Because stimulated emission contributes an insignificant portion of the solid-angle-integrated emergent intensity, natural planetary lasers have no real effect on their planets' atmospheric radiative equilibrium. However, a long-path single-pass gain of 7 % normal to the subsolar zenith, being comparable to single-pass gain in laboratory lasers, is nominally large enough to overcome reflection losses if a resonator were configured to use it [Deming & Mumma, 83] [Mumma, 83]. Gordiyets and Panchenko [82] project a possible laser power of 360 erg/(cm².sec) for 3 % mirror losses. Thus natural planetary lasers provide ready-made gain media of enormous size, possibly yielding, if engineered, useful lasers of high specific power.

Controlling Light

The most powerful laser imaginable is useless if its light cannot be controlled. A planetary laser capable of communicating over interstellar distances must exercise three types of laser control: tilt, modulation and phase. Tilt control means directing the laser, both the intracavity beam and the coupled output beam. Steering the beams produced by a large planetary laser is a spacecraft system problem treated by Part 2. Modulation is the willful distortion of the output

beam that would otherwise propagate as a plane wave, with the goal of impressing it to carry information. Chapter 3 outlines salient theory of modulating optical carriers, while again Part 2 examines ways of performing it practically for the planetary laser. Phase control, more subtle yet fundamental, insures that we have a plane wave (or at least a known, predictable waveform) to modulate. More basically, it permits the laser cavity to develop a resonant field in the first place. Without phase control, there can be no laser.

As a simple illustration of what phase control does, imagine a plane wave reflecting off a mirror surface. In general, the reflected wave will be phase-distorted in addition to being redirected. That is, any departure of the reflecting surface from flatness causes some portions of the wave to be retarded compared to others, which if excessive ruins the coherence of a laser beam. Clearly the oscillating field inside a laser cavity will degenerate if with each reflection it accumulates more phase distortions; once the field is no longer self-resonant, lasing stops. And a propagating wavefront with disturbed phase will interfere with itself, resulting in diminished far-field intensity. When its far field is thousands of km away (the planetary laser intracavity distance), or dozens of pc away (the interstellar transmission distance), a laser cannot afford uncompensated phase inaccuracies if it is to perform well. Rather, its optical quality should as much as possible be diffraction limited, compromised by the physical nature of light itself rather than by imperfections in the mechanisms which control the light.

Phase distortions also occur upon transmission through materials, as for example lenses or fluid media. Fixed distortions which result from manufacturing inaccuracies and material defects, such as variations in surface figure,

assembly alignment or material composition can generally be reduced with higher costs. More difficult to compensate are time-varying and often unpredictable distortions arising from thermal changes, jitter, and creep in solid media, or turbulence and other transient inhomogeneities in fluids. The classic problem sources for laser transmission using spacecraft, all potential problems for engineering planetary lasers, are respectively: differential expansion from the steep thermal gradients encountered by moving in and out of sunlight, undamped vibrations from momentum transfer elsewhere in the structure, and laser propagation through an atmosphere.

The field of adaptive optics arose over the past few decades to address particularly these time-varying problems [Pearson, 79]. In fact, Pearson regards some form of adaptive optics as "essential" for space-based, large aperture optical devices. Although system details vary, the three basic parts of an adaptive optical train can be represented by an optical wavefront sensor, a closed-loop control network, and a mechanically deformable optical element. The Coherent Optical Adaptive Technique (COAT), for instance, works by monitoring an outgoing wave, calculating error signals based on comparing these sensory data to the desired waveform, generating control signals, and then adjusting deformable optics to reduce the error. Obviously, hardware complexity and mass, as well as the system bandwidth required, would depend directly on the local operating environment, expected timescale of phase variations needing compensation, and optical wavelength used.

The typical spacecraft operating environment is extreme in terms of physical speed, thermal variation, lack of viscous damping, lack of a reaction mass "sink", and because of overwhelming requirements for robustness, fault-tolerance and reliability, given the difficulty of repair. We could expect phase distortions to occur with time periods ranging all the

way from years (for orbital variations) to hours (for attitude variations) to seconds (for structural dynamic waves) down to milliseconds (for local structural vibrations and gain medium turbulence in the planetary mesosphere). Controlling optical element figure error for systems operating at a wavelength of $10.6\text{ }\mu\text{m}$ bridges the gap between being fairly simple (as it is for radio wavelengths) and virtually impossible (as it is for visible wavelengths). The standard criterion of $\lambda/20$ surface accuracy for diffraction limited performance is then about $0.5\text{ }\mu\text{m}$, feasible either monolithically or with numerous actuators and a flexible mirror (or some combination of both, as we outline in Part 2). Pearson proposes of order 100 actuators to achieve arbitrary figure control at $10\text{ }\mu\text{m}$ for a 1 m diameter flexible mirror. An actual device built by Stephens and Lind [78] performed well at $10\text{ }\mu\text{m}$ using a hexagonal close-packed actuator array with 2 cm spacing. Clearly, for larger components, the number of actuators required by this type of control could become huge, with the control system correspondingly complex.

Other problems plague continuously flexible mirrors for use in space. First, most employ membrane mirrors because they are easily deformed with small applied forces. Such surfaces can be metallized polymers [Chown, 85] or, more appropriately for the materials-degrading space environment, vapor-deposited metal membranes about $1\text{ }\mu\text{m}$ thick, of Ti or its alloys, Ni, Be, or Mo [Grosso & Yellin, 77]. Deformation is typically accomplished with electrostatic fields, applied either discretely by electrodes or continuously by scanning charged-particle guns. Such a device obviously requires the strictest of environmental field control for proper functioning, something not intrinsically guaranteed by interplanetary space (isolating an optical element by enclosing it within an opaque metal Faraday cage would appear self-defeating). Second, resolving the ambiguity of surface

errors which produce phase shifts in even multiples of 2π most likely necessitates using two different measurement wavelengths [Palma et al, 79]. Finally, monitoring the far-field is not even possible in our case; we would be constrained to infer far-field properties by measuring near-field properties. It is unlikely that such indirect sensing would yield far-field accuracy to the diffraction limit, and quite likely that the tapping device itself would distort the outgoing beam yet more. The sensitive nature of these techniques, as well as their physical complexity, naturally reduces their attractiveness even to a designer who has no other choices.

As early as 1978, Stephens and Lind recognized the advantages of intracavity compensation for lasers. Presuming a COAT scheme, they showed that correcting phase distortions inside the resonator cavity "before they can diffract into intensity variations" maximizes far-field intensity, and requires smaller corrector surface excursions (thus a lighter mechanism) than extracavity post facto correction. In any case, we already expect that intracavity phase control is needed to make a planetary laser oscillate in the first place. If the right waveform is programmed onto the COAT optics, a reflected wave, pre-compensated for the distortions it will incur on its return pass back through the resonator, can "be coupled out of the resonator as a diffraction-limited beam." Stephens and Lind [78] used an 18-element intracavity mirror to correct satisfactorily the effects of resonator misalignment, mirror figure errors, laser medium inhomogeneities, and extracavity optical train imperfections for a CO₂ laser.

That kind of fine control applied to a mirror with diameter of order just 5 m would require thousands of deformation zones, however; assembling a large system with, in

turn, hundreds of thousands of such mirrors would not bode well for reliable operation in the harsh, remote environment of interplanetary space. A better system, if such microscopically fine adaptability proves necessary, might be the all-optical self-referenced Interference Phase Loop (IPL) [Fisher, 85], probably the acme of current mechanical phase compensation systems. The core device is a monolithic optically-addressed Photo-Emitter Membrane Light Modulator (PEMLM) consisting of three parts. At the back is a photocathode, upon which is projected (from behind) the error-proportional output intensity pattern from an optical phase sensor. The subsequent electron image is amplified directly by the middle element, a Micro Channel Plate (MCP) array of tiny multiplier pores lined with semiconducting glass, to deflect the 10 μm diameter deformable membranes covering each pore. Thus the front mirror element (the membrane surfaces), being essentially continuously spatially modulated by the conjugate of an input wave, effects phase and optical path distortion compensation "over a multi-wave dynamic range with no $2n\pi$ or phase quadrant ambiguity."

The IPL's behavior is self-centering and therefore stable, immune to wave amplitude fluctuations, and operative on even partially coherent and multispectral light. Its robust monolithic configuration can easily be extended to resolutions of more than a million elements, and can be operated as a "high-resolution bistable/multistable element in optical information processing applications", meaning that it can be used as an optical switch. Thus devices based on the IPL principle could achieve laser cavity mode control, beam microsteering, precise adaptive figure control, or signal modulation.

Nonlinear Optical Phase Conjugation

An alternative approach to controlling simultaneously the direction, phase and modulation of light results directly from the physics of electromagnetic radiation propagating in materials. The most recent and startling branch of this field, called Nonlinear Optical Phase Conjugation (NOPC), "involves the real-time spatial and/or temporal information processing of electromagnetic fields" [Pepper, 82]; it has been the target of intensive study for merely a decade, being one of the many areas opened only by the advent of lasers. Anticipated applications range from advanced spectroscopy, interferometry and ultralow noise detection, to optical computing, image processing, optical signal processing in both the time and spatial domains, and real-time adaptive optics. When feasible, the all-optical NOPC techniques can sidestep such elaborate engineering stunts as COAT; NOPC replaces "cumbersome, costly" electromechanical components and often performs with better spatial and temporal bandwidths [Pepper, 82]. Lind et al [81] found that under ideal conditions NOPC phase compensation quality is over two orders of magnitude better than that of contemporary deformable mirrors.

A material's field-dependent susceptibility χ can be written [Yariv & Fisher, 83] as a power-series expansion in terms of the total electromagnetic field E :

$$\chi(E) = \chi^{(1)} + \chi^{(2)}E + \chi^{(3)}E^2 + \dots \quad (2.1)$$

such that the material's polarization $P = E\chi$ is simply:

$$P(E) = E\chi(E) = \chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \dots \quad (2.2)$$

E is composed in general of many waves with different frequencies, polarizations, and k vectors (propagation directions). The first, second and third order terms in the polarization expansion describe recognizable interactions of light in matter.

The linear, $\chi^{(1)}$ terms describe well-known linear optical properties such as absorption, gain, index of refraction, and birefringence, which couple a light wave with matter to produce a "second" wave which is either attenuated, amplified, or redirected. In a material exhibiting only linear susceptibility terms, separate waves pass through each other without interaction, as though each were there alone. Saturation of these linear effects, and many other important behaviors, must be described by higher-order terms.

The second-order, $\chi^{(2)}$ terms produce second order (nonlinear) interactions. Since the $\chi^{(2)}E^2$ polarization factor will contain a cross-term, it can be used to identify an interaction between two separate input waves and the material. The newly radiating polarization comprises a third, generated wave; thus second-order interactions are called three-wave mixing. Occurring only in materials lacking inversion symmetry, they are: second-harmonic generation (in which the new wave has double the frequency of identical input waves), optical rectification (in which the input waves' time-varying components cancel out), parametric mixing (in which the new wave frequency is the sum of the input frequencies), and the Pockels effect (changes in refractive index due to induced electrostatic fields).

The third-order, $\chi^{(3)}$ terms will obviously lead to an expanded polarization factor containing yet more cross-terms, thereby describing interactions of three input waves and the material to produce a new, fourth light wave. Such four-wave

mixing occurs, regardless of inversion symmetry, in an enormous variety of materials in all states, requiring only that the medium exhibit a "large" $\chi^{(3)}$ susceptibility. Important effects not relevant to this study include: two-photon absorption, the dc Kerr effect, dc-induced harmonic generation (the third-order contribution to second-order harmonic generation) and third-harmonic generation (analogous to second-harmonic generation). The other major $\chi^{(3)}$ effects can all be used for NOPC. They are: Stimulated Raman Scattering (SRS, in which light scatters from molecular vibrations in solids), Stimulated Brillouin Scattering (SBS, in which light scatters from sound waves propagating through a fluid, discussed in Appendix 2-1), Nondegenerate Four-Wave Mixing (NFWM), and Degenerate Four-Wave Mixing (DFWM, the instantaneous ac Kerr effect). These last two, discussed in Appendix 2-2, involve light waves scattering off the phase gratings caused by their mutual interference in transparent media for which the speed of light "depends linearly upon the light intensity" [Yariv & Fisher, 83]. SBS and DFWM are the preferred, and most-studied, methods for phase conjugation.

The expanded polarization form of equation 2.2 clearly shows how these subtle effects are directly related through fundamental physics to more familiar phenomena like index of refraction. Materials exhibiting absorption and gain features, called resonant media, cannot be analyzed quite so easily, because terms of arbitrarily high order in the susceptibility expansion can become significant; nonetheless several experiments in NOPC for resonant media have been carried out. We might expect that using this physics of light interactions directly could enable us not only to perform simple functions like steering and modulating (by redirecting and switching the light), but also to control its phase properties as well. Indeed, the phase compensation abilities of NOPC are its most exciting prospect.

The light wave generated by the third order susceptibility effect of an ideal NOPC technique is the phase conjugate of the input probe wave. That means it has the same frequency ω but a spatial complex amplitude which is the complex conjugate of the input amplitude [Pepper, 82]. The phase conjugate wave behaves like a time-reversed replica of the probe wave. That is, the nonlinear interaction exactly reverses both the direction and phase of the incoming light, so that the new wave generated by the interaction propagates "backwards", its wave fronts coinciding everywhere with those of the probe. Put simply, "the conjugate field can be viewed as equivalent to the incident field traveling backward in time" [Giuliano et al, 79].

A device which can produce phase conjugate replicas for input waves of some particular frequency can be referred to as a Phase Conjugate Mirror (PCM). To appreciate the profound implications of NOPC for optics, we consider the now classic "thought demonstrations" illustrating PCM behavior. An ideal lossless PCM reverses the direction, polarization and phase of incoming monochromatic light; it "reverses all the quantum numbers of the incident photon" [Pepper, 82]. Thus no linear or angular momentum transfer can occur between the photons and the PCM, so the mirror feels no radiation pressure or torque. Because the \mathbf{k} vector is exactly reversed, light striking a PCM retraces its path back to the source, regardless of the PCM's tilt. Thus a diverging wave coming in will be reflected as a converging wave going back out, refocusing on its source.

Looking into a PCM, an observer would see nothing, because the only light striking the eyes would be that which had emanated from the eyes (that is, not the rest of the viewer's face) [Yariv & Fisher, 83]. Further, because of the light's

phase reversal, transparent objects between the viewer and the PCM would appear invisible in it, since phase distortions imprinted on the transmitted wave, which normally make such an object visible, are removed from the conjugate wave on its backward pass [Gower, 84].

How can such a device be used? Since any aberrator will remove from the conjugate wave all the distortions it introduced into the probe wave, as long as all the aberrated light enters the PCM, any double-pass optical train incorporating a PCM can transmit diffraction-limited light no matter how poor its optics are and despite unpredictable transient or permanent variations in the optical path. The PCM does this passively and without massive and complex equipment.

If an object is illuminated such that a reflected glint passes through a high-power amplifier to strike a PCM, the conjugate wave, predistorted to compensate for imperfections in its second pass back through the amplifier, optical train, and intervening atmosphere, will return to the object, feasibly with enough energy to destroy it. This automatic pointing, tracking and targeting feature is envisioned to be useful for laser fusion devices (we should clarify incidentally that since such a system would destroy without discrimination anything it looked at, applications to SDI could only be Strangelovian).

All-optical information processors for computing and data transmission, if based on NOPC, would realize many advantages over conventional machines: "enhanced spatial and temporal bandwidths, reduced size, cost, weight, and power consumption, and improved environmental resistance to RFI [Radio Frequency Interference], vibration, and temperature" [O'Meara et al, 83]. Spatial properties of NOPC can affect focusing, imaging, transmission, and generation of monochromatic light, with applications to: imaging through optical fibers, lensless

photolithography, speckle-free imaging, pattern recognition, arithmetic image processing, parallel logical processing, holographic storage, edge enhancement and interferometry, in addition to the aberration compensation and tracking uses already described. Temporal properties of NOPC can be important for encoding information, optical filtering, pulse compression, pulse sequence reversal, and computing.

One critical application, particularly relevant to this study, uses both spatial and temporal properties by employing PCMs to make laser resonators, or Phase Conjugate Resonators (PCRs). With one of the conventional mirrors of a laser oscillator replaced by a PCM, both static and dynamic intracavity distortions can be reduced, resulting in "a diffraction-limited output from the conventional mirror end of the resonator, subject to the precision fabrication of the mirror", thus "[breathing] new life into systems that in the past were deemed impractical because of reflection losses" [O'Meara et al, 83]. Because any electromagnetic field will reproduce itself after two round trips in a PCR, this kind of resonator is stable no matter what curvature the regular mirror has. A PCR can "oscillate satisfactorily in the face of aberrations sufficiently severe to quench the oscillation of a matching conventional resonator" [O'Meara et al, 83]. One such resonator functioned even with a kitchen spatula as the regular mirror [Feinberg, 83]! Most PCRs use FWM as the nonlinear interaction; because the energy introduced by several input waves lets the PCM operate with gain, in many cases no gain medium is even needed in the mirror cavity for oscillation to occur. Alternatively, a single input wave can by multiple reflections itself yield the other mixing waves, so that the PCR is self-pumped. All such systems must have a startup feature, though, since initially there is no reflectivity at the PCM.

No NOPC technique can perform all these miracles, and no NOPC technique is ideal even for its optimal applications. All the methods have limitations and restrictions, detailed for the most eligible candidates, SBS and FWM, in Appendices A2-1 and A2-2 respectively. Still, although the field is yet nascent, the techniques are real. Primarily a research tool so far, NOPC has nonetheless:

"been observed in a myriad of states of matter (solids, including semiconductors; liquids; gases and vapors; liquid crystals; aerosols; and plasmas), using a variety of... interactions..., employing lasers that span the optical spectrum (from the UV to the IR), and using pulsed and cw lasers (from megawatts to microwatts). The response times of the optical nonlinearities range from seconds to picoseconds" [Pepper, 82].

Clearly the potential uses of NOPC are far-ranging. Particularly given the tremendous progress attained in only its first decade of study, we can project the dependence of many future devices on its physics. Beyond the lure of its apparent ability to solve some of the worst problems of space-based, large aperture, high-power lasers, NOPC's certain role in advanced optical systems means that it must be addressed as a possible phase control tool, along with the mechanical methods cited earlier, by any study of planetary lasers.

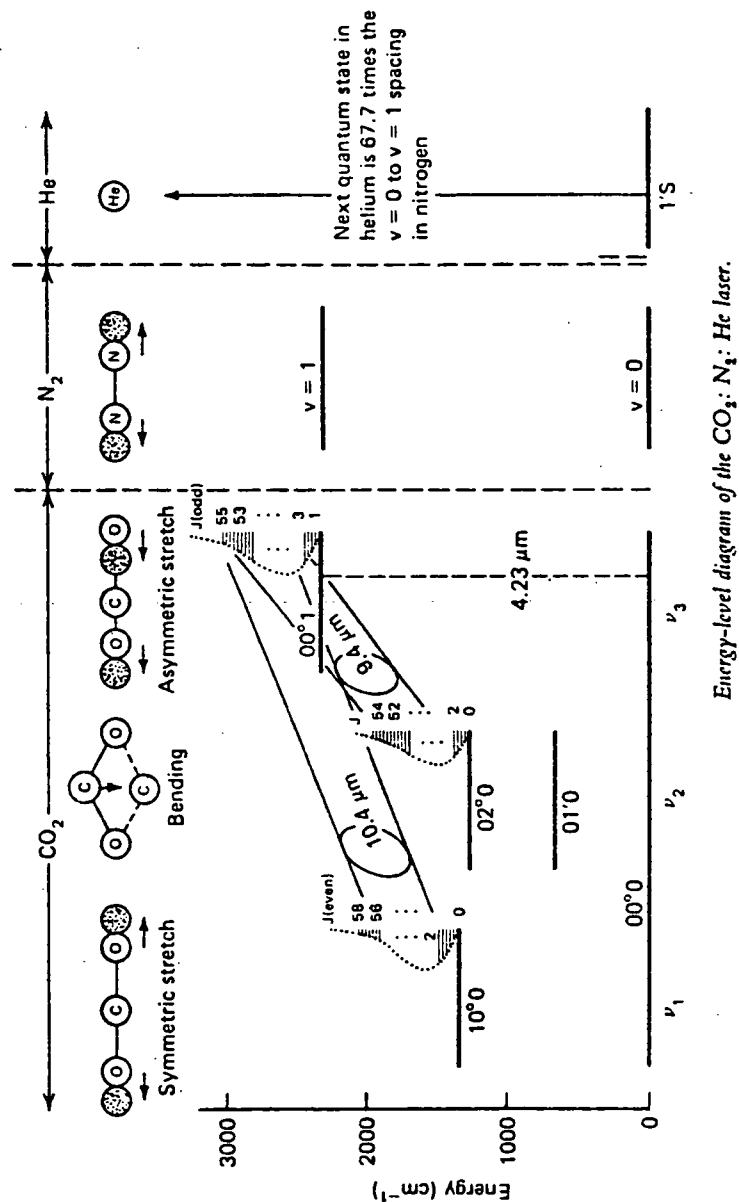


Figure 2-1 Vibration modes and energy level diagram for a CO₂ laser system. [Verdeyen, 81]

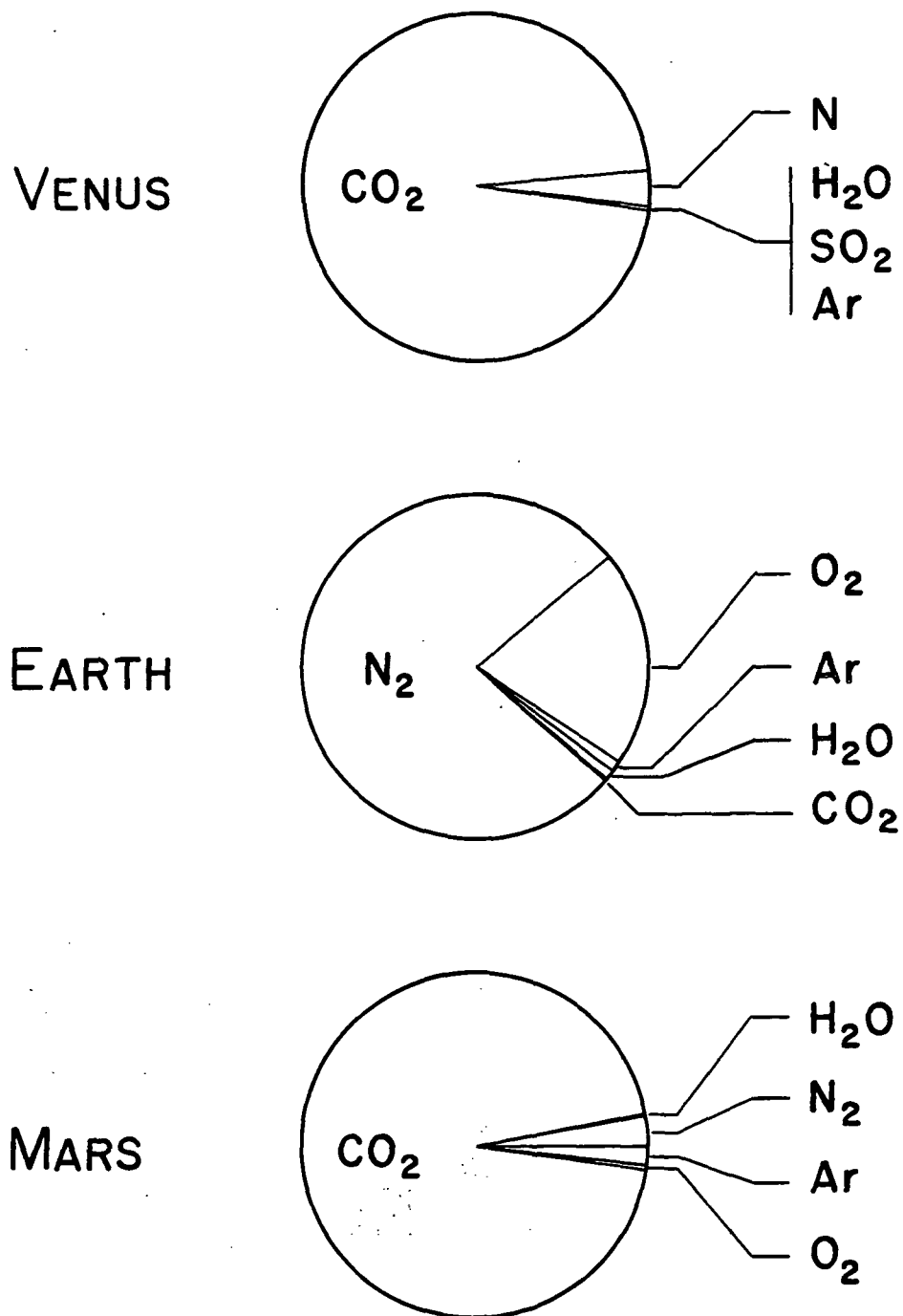


Figure 2-2 The major terrestrial planet atmosphere constituents.
after [Beatty et al, 82]

Appendix A2-1 Stimulated Brillouin Scattering.

Both Stimulated Raman Scattering (SRS) and Stimulated Brillouin Scattering (SBS) involve interactions between light waves and non-electromagnetic waves propagating in materials. SRS scatters light waves from molecular vibrations in a solid as its moving molecules change the solid's polarizability. The new, scattered light wave is downshifted in frequency by an amount equal to the molecular vibration frequency; albeit independent of scattering angle, this shift is large, typically 500 to several thousand cm^{-1} . IR radiation around $10.6 \mu\text{m}$ measures about 950 cm^{-1} , so SRS would at least double the light's wavelength, precluding its use for a PCR.

SBS also is based on inelastic photon scattering [Pepper, 82], because the quantum state of its medium changes in the interaction. It occurs in materials, such as many fluids subject to the electrostrictive effect, where polarizability is a function of pressure. In these media, light can be scattered by pressure, or density, waves or acoustic phonons. The coupling works as follows:

"...a strong electric field can be produced by the passage of an intense light beam. Through electrostriction, this results in periodic changes in the density of the medium and therefore in the medium index, which generates a traveling acoustic wave which, in turn, scatters (reflects) some of the input optical beam" [Giuliano et al, 79].

Tightly feedback, this mechanism when operating above its threshold can convert a large fraction of the input wave to the

scattered wave, and "the fraction of nonconjugated return is immeasurably small" [Giuliano et al, 79].

The backscattered, conjugate wave arises from spontaneous noise, and grows preferentially with a gain about twice that of competing modes. Because the Bragg (interference) mirror set up in the medium moves at the speed of sound, the scattered wave is Doppler shifted by the frequency Ω of the sound wave down from the input frequency ω . Its direction is also altered from the input wave vector \mathbf{k} by the sound wave's vector \mathbf{q} . The frequency shift is angle-dependent; for backward scattering (most often used for NOPC) that shift is "twice the refractive index times the ratio of the material's speed of sound to the vacuum speed of light" and therefore around one part in 10^4 or 10^5 [Yariv & Fisher, 83].

The phase compensation ability of SBS was recognized and demonstrated as early as 1972 [Nosach et al, 72] [Zel'dovich et al, 72] but not theoretically explained until years later. Generally regarded as the simplest, and perhaps most efficient, of the NOPC methods, it requires only one input wave (the probe) because it generates its own mixing waves, and can occur in practically any solid, liquid, or gaseous medium as long as "the coherence time $1/\Delta\omega$ of the incident light is long compared to the response time τ of the acoustic phonons" [Hon, 82]. Since τ is of order 10^{-8} seconds even in gases, this coherence time requirement is generally not difficult to meet.

Although SBS efficiencies (the ratio of backward-going energy to incident energy) have been measured as high as 90 %, the range 30 - 50 % is used for system design. The quality of Wave Front Reversal (WFR) is nearly perfect if the Fresnel number (see Figure 7-7) is small (< 1) and the SBS occurs in a light pipe [Hon, 82], especially one with rough internal

surfaces to promote mode mixing (the light guide increases interaction length and hence the percentage of backscatter without increasing the power density so much that the SBS medium breaks down [Giuliano et al, 79]). Thus its simplicity, efficiency, quality, and its ability to use large interaction volumes make SBS ideal for high-power NOPC uses.

Liquid CS₂, ether, acetone, CCl₄, plasmas, and compressed SF₆ (18 atm) and CH₄ (125 atm) have been used as SBS media. Major limitations are that it is a threshold phenomenon, does not conjugate the polarization state of the backward-going wave and requires a uniformly polarized probe wave, would require an enormous pressure-containing light pipe for really large applications, seems limited by background noise to resolutions of order 10⁷ pixels in the steady state [Hellwarth, 83b], and with repeated passes walks off the laser gain profile because of the cumulative Stokes frequency downshift [Giuliano et al, 79]. That cumulative frequency shift in particular constitutes a rather intractable problem for intracavity use in a PCR.

Appendix A2-2 Four Wave Mixing.

Four Wave Mixing (FWM), first proposed for NOPC by Hellwarth in 1977, is an optical parametric interaction based on elastic photon scattering [Pepper, 82], meaning that it leaves the nonlinear medium in the same quantum state after the interaction as beforehand. The canonical geometry consists first of two strong, counterpropagating pump waves of equal intensity and with the same distortions (that is, phase conjugates [Feinberg, 83]), impinging on the nonlinear mixing medium. A third, probe wave is introduced at some arbitrary angle to the pump line; FWM generates a fourth wave, proportional to the probe's spatial complex conjugate, which then propagates back along the probe's direction. If the probe and pumps are all at the same frequency ω , Degenerate Four Wave Mixing, the most studied kind, occurs. Any other condition is called Non-Degenerate FWM, which produces interesting results shown later.

The ability of a material such as a solid photorefractive crystal, having $\chi^{(3)} > 10^{-2} \text{ cm}^3/\text{erg}$ [Dunning et al, 84]), to couple waves by interference scattering arises through an elaborate sequence of events [Feinberg, 83]: charges in the crystal structure migrate when illuminated, producing strong local electrostatic fields which then change the medium's refractive index through the Pockels effect. The steady state refractive changes depend only on relative intensities in the interaction volume, although the speed of these changes increases with increasing input intensity, down to response times of order less than nsec [Feinberg, 83].

Although a quite distinct process, FWM is often explained by analogy as operationally equivalent to real-time holography.

That is, the probe wave E_p and pump wave E_1 interfere to generate a grating, which is read out by pump wave E_2 to produce part of the conjugate wave E_c . Another, coherently superposed part of E_c , is read out by E_1 from the grating formed by the interference of E_p with E_2 [Pepper, 82].

To ensure pump matching, most laboratory setups use a beam splitter to derive both pump waves from one laser source. If the medium's linear losses are low, the second pump can be made simply by retroreflecting the first pump after passage through the mixing volume. Novel, and experimentally successful variations include mixing in long, narrow optical waveguides such as fibers (which reduces by several orders of magnitude the required pump power while increasing by several orders of magnitude the number of resolution elements attainable [Hellwarth, 83a]), and using the strong counterpropagating circulating field inside a laser oscillator to constitute the pumps (since the counterpropagating fields are already aligned, and since the intracavity field strength is much stronger than the outcoupled field strength, greater nonlinear gains can be achieved) [Pepper & Yariv, 83]. FWM has been observed in both internal [Feinberg, 82] and external [Giuliano et al, 79] self-pumping arrangements (in which a separate startup pump is used to initiate NOPC oscillation until feedback pumping can take over), and in unidirectional [White et al, 1982] and bidirectional [Lind et al, 81] ring laser geometries.

Since its behavior most closely approximates that of an ideal PCM, DFWM is normally considered for PCR uses. Because of the double round trip stability inherent in PCRs, normal $c/2L$ longitudinal cavity modes (where $L \equiv$ the resonator length) are replaced in them by paired half-axial modes, spaced at $\pm c/4L$ and centered about the pump frequency; the PCR oscillation frequency is locked to that of the pumps, which by

virtue of their power input also control the gain value of the PC mirror.

One special, flexible feature of FWM may prove useful for space-based distributed optical systems. Consider even the most basic resonator configured to oscillate a mesospheric laser, consisting of two mirror satellites orbiting the planet and defining a tangential path through the atmospheric layer which lases whenever subsolar. Two immediate problems arise. First, the frequency of light leaving each mirror is oppositely Doppler-shifted relative to the gain profile of the atmosphere for any orbit other than synchronous, producing a two-color laser. Also, both mirrors move at orbital velocities during the time it takes the light to travel in a straight line between them, so mirror pointing cannot be normal. However, using FWM we could purposely misalign the pump angles and/or frequencies, called pump detuning, thereby simultaneously achieving both pointing offset and Doppler frequency compensation, still with perfect phase matching [Giuliano et al, 79].

Although pump misalignment reduces efficiency, this negative aspect is minimized for small probe/pump angles. In fact, for any independent combination of modest point-ahead angle and frequency shift, there exists one three-dimensional (non-coplanar) geometry of pumps and probe which will yield perfect phase matching [Lind et al, 81]. Additionally, any spatial or temporal perturbations impressed on the pump will be transferred (to first order) directly to the conjugate wave, opening possibilities for signal modulation [Pepper, 82]. Here then is simultaneous tilt, modulation, phase and frequency control of light.

Because of symmetry considerations any material can be used for FWM [Pepper, 82], but a large $\chi^{(3)}$ is desirable for

efficient operation. FWM has been performed in photorefractive crystals, semiconductors, glasses, organic compounds, liquid dyes and liquid crystals, atomic vapors and gases [Pepper, 82]. Work is under way to develop long-chain organic molecules with "giant values of susceptibility" due to their strongly overlapping π bonds [Dunning et al, 84]. Because of the pumps' energy contribution, reflectivities $\gg 1$ are possible. Including pump depletion though, overall efficiencies are $\leq 20\%$ for conjugate energy reflectivities of 50% , the optimal value for equal intensities of probe and pumps. Large conjugate reflection coefficients require large absorption coefficients; for use in an inverted medium, low intensity operation close to line center is best [Lind et al, 81].

FWM has been used extensively to conjugate CO₂ laser light at $10.6\ \mu\text{m}$ in several media. As long ago as 1982, pulsed $10.6\ \mu\text{m}$ light had been conjugated in the solids HgCdTe, KCl:ReO₄, and Ge (at 800% reflectivity!), and in the gases SF₆ and inverted CO₂ (both of which are resonant media). CW $10.6\ \mu\text{m}$ light had been conjugated in HgCdTe and SF₆ by the same time, as well as simultaneous multi-line CO₂ laser light.

Restrictions, of course, abound: for efficient conjugation, the source laser should operate in single longitudinal and transverse modes. In a non-guided geometry, the pump waves must be nearly planar. Conjugation efficiency is improved for larger interaction volumes, practically attainable only with a collinear pump/probe configuration [Pepper & Yariv, 83]; on the other hand, collinear or "small-angle" geometries can "wash out" one of the scattering gratings if the relaxation time of the medium is of the same order as the time constant of its molecular motion, thus degrading reflectivity [Lind et al, 81]. Moreover, any path aberrations must be effectively imaged at the interaction volume (just as with SBS all the diffracted light must enter the medium) if they are to be conjugated. The

frequency and directional detuning abilities of FWM, while versatile, nonetheless require exacting geometries, and FWM remains untested at large transverse scales.

References

John Auyeung, D Fekete, D M Pepper, A Yariv- "A Theoretical and Experimental Investigation of the Modes of Optical Resonators with Phase-Conjugate Mirrors" IEEE Journal of Quantum Electronics Vol QE-15 No 10 (Oct 1979).

J K Beatty, B O'Leary, A Chaikin (eds)- The New Solar System 2nd ed (Cambridge, 1982).

Ernest E Bergmann, Irving J Bigio, B J Feldman, Robert A Fisher- "High-Efficiency Pulsed 10.6- μ m Phase-Conjugate Reflection via Degenerate Four-Wave Mixing" Optics Letters Vol 3 No 3 (Sep 1978).

Marcus Chown- "Doubts Surround Strathclyde's Mirror for Space Weapons" New Scientist (5 Sep 1985).

Mark Cronin-Golomb, Baruch Fischer, Joseph Nilsen, Jeffrey O White, Amnon Yariv- "Laser with Dynamic Holographic Intracavity Distortion Correction Capability" Appl. Phys. Lett. Vol 41 No 3 (1 Aug 1982).

Mark Cronin-Golomb, Baruch Fischer, Jeffrey O White, Amnon Yariv- "Passive (Self-Pumped) Phase Conjugate Mirror: Theoretical and Experimental Investigation" Appl. Phys. Lett. Vol 41 No 8 (15 Oct 1982).

M J Damzen & Henry Hutchinson- "Laser Pulse Compression by Stimulated Brillouin Scattering in Tapered Waveguides" IEEE Journal of Quantum Electronics Vol QE-19 No 1 (Jan 1983).

D Deming, F Espenak, D Jennings, T Kostiuik, M Mumma, D Zipoy- "Observations of the 10 μ m Natural Laser Emission from the

Mesospheres of Mars and Venus" Icarus 55, 347-355
(1983).

Drake Deming & Michael J Mumma- "Modeling of the 10- μ m
Natural Laser Emission from the Mesospheres of Mars and
Venus" Icarus 55, 356-368 (1983).

R M Dickinson & J T English- "Retrodirective - Optical -
Transponder Concept" NASA Tech Briefs Vol 9 No 2
(Summer 1985).

G J Dunning, M B Klein, R C Lind- Phase Conjugate Optics Final
Technical Report N84-30951 (Feb 1984).

Jack Feinberg- "Self-Pumped, Continuous-Wave Phase Conjugator
Using Internal Reflection" Optics Letters Vol 7 No 10
(Oct 1982).

- "Optical Phase Conjugation in Photorefractive
Materials" in: [Fisher, 83].

Arthur D Fisher- "Self-Referenced High Resolution Adaptive
Wavefront Estimation and Compensation" SPIE Proc.
Vol 551 Adaptive Optics (1985).

Robert A Fisher (ed)- Optical Phase Conjugation
(Academic Press, 1983).

R H Freeman & J E Pearson- "Deformable Mirrors for All
Seasons and Reasons" Applied Optics Vol 21 No 4
(15 Feb 1982).

C A Giuliano, R W Hellwarth, R K Jain, R C Lind, T R O'Meara, G M Wandauro, V Wang- Correction of Phase Distortion by Nonlinear Optical Techniques Interim Technical Report Hughes Contract N00014-77-C-0593 (Mar 1979).

B F Gordiyets & V Ya Panchenko- "Infrared Radiation and Inversion Population of CO₂ Laser Levels in Venusian and Martian Atmospheres" translated from Russian: P N Lebedev Institute of Physics, USSR Academy of Sciences (1982) NASA TM-77076 (1983).

M C Gower- "The Physics of Phase Conjugate Mirrors" Prog Quant Electr Vol 9 pp 101-147 (1984).

Ronald P Grosso & Martin Yellin- "The Membrane Mirror as an Adaptive Optical Element" J. Optical Soc. America Vol 67 No 3 (Mar 77).

Marquet & Hardy- "Historical Review of Adaptive Optics Technology" SPIE Proc. Vol 141 Adaptive Optical Components (1978).

William K Hartmann- Moons and Planets 2nd ed (Wadsworth, 1983).

R W Hellwarth- (a) "Phase Conjugation by Four-Wave Mixing in a Waveguide" in: [Fisher, 83].

- (b) "Phase Conjugation by Stimulated Backscattering" in: [Fisher, 83].

David T Hon- "Applications of Wavefront Reversal by Stimulated Brillouin Scattering" Optical Engineering Vol 21 No 2 (Mar/Apr 1982).

Ivan P Kaminov & Anthony E Siegman- Laser Devices and Applications (IEEE, 1973).

R C Lind, W B Brown, C R Giuliano, R K Jain, J F Lam, B M Merchant, R A McFlarlane, T R O'Meara, G C Valley-
Correction of Phase Distortion by Nonlinear Optical Techniques Final Technical Report N81-28425
(May 1981).

R C Lind & D G Steel- "Demonstration of the Longitudinal Modes and Aberration-Correction Properties of a Continuous-Wave Dye Laser With a Phase-Conjugate Mirror"
Optics Letters Vol 6 No 11 (Nov 1981).

R C Lind, D G Steel, M B Klein, R L Abrams, C R Giuliano, R K Jain- "Phase Conjugation at 10.6 μ m by Resonantly Enhanced Degenerate Four-Wave Mixing" Appl. Phys. Lett. Vol 34 No 7 (Apr 1979).

Michael J Mumma- personal memorandum to: James R Bogan, Jet Propulsion Laboratory (23 Jul 1983).

Michael J Mumma, David Buhl, Gordon Chin, Drake Deming, Fred Espenak, Theodor Kostiuik, David Zipoy- "Discovery of Natural Gain Amplification in the 10-Micrometer Carbon Dioxide Laser Bands on Mars: A Natural Laser" Science Vol 212 pp 45-49 (3 Apr 1981).

O Yu Nosach, V I Popovichev, V V Ragul'skii, R S Faizullov- "Cancellation of Phase Distortions in an Amplifying Medium with a 'Brillouin Mirror'" translated in JETP Letters Vol 16 p 435 (1972).

Thomas R O'Meara, D M Pepper, Jeffrey O White- "Applications of Nonlinear Optical Phase Conjugation" in: [Fisher, 83].

- Gary E Palma, Robert K Elkow, Albert W Angelbeck-
 "Two-Wavelength Phase Control" SPIE Proc. Vol 179
Adaptive Optical Components II (1979).
- C K N Patel- "High-Power Carbon Dioxide Lasers" Scientific
 American Vol 219 (Aug 1968).
- James E Pearson- "The Whither and Whether of Adaptive Optics"
SPIE Proc. Vol 179 Adaptive Optical Components II
 (1979).
- David M Pepper- "Nonlinear Optical Phase Conjugation" Optical
 Engineering Vol 21 No 2 (Mar/Apr 1982).
- David M Pepper & Amnon Yariv- "Optical Phase Conjugation
 Using Three-Wave and Four-Wave Mixing via Elastic Photon
 Scattering in Transparent Media" in: [Fisher, 83].
- Anthony E Siegman- [op cit Chapter 7].
- A E Siegman, Perre A Belanger, Amos Hardy- "Optical Resonators
 Using Phase-Conjugate Mirrors" in: [Fisher, 83].
- Duncan G Steel and Juan F Lam- "Multiline Phase Conjugation in
 Resonant Materials" Optics Letters Vol 5 No 7
 (Jul 1980).
- G I Stepanova & G M Shved- "The Natural 10- μ CO₂ Laser in
 the Atmospheres of Mars and Venus" Soviet Astronomy
 Letters 11(3) (May-June 1985).
- Ronald R Stephens & Richard C Lind- "Experimental Studies of
 Adaptive Laser Resonator Techniques (ALERT)" SPIE Proc.
 Vol 141 Adaptive Optical Components (1978).

- "Experimental Study of
an Adaptive-Laser Resonator" Optics Letters Vol 3 No 3
(Sep 78).

B R Suydam & Robert A Fisher- "Transient Response of
Kerr-Like Phase Conjugators" in: [Fisher, 83].

H Vanherzeele, J L Van Eck, A E Siegman- "Mode-Locked Laser
Oscillation Using Self-Pumped Phase-Conjugate Reflection"
Optics Letters Vol 6 No 10 (Oct 1981).

Joseph T Verdeyen- Laser Electronics (Prentice Hall, 1981).

Hanns J Wetzstein- "A Tutorial Review of Key System Aspects of
Wavefront Measurement, Control and Conjugation" SPIE
Proc. Vol 141 Adaptive Optical Components (1978).

Jeffrey O White, Mark Cronin-Golomb, Baruch Fischer, Amnon
Yariv- "Coherent Oscillation by Self-Induced Gratings in
the Photorefractive Crystal BaTiO₃" Appl. Phys. Lett.
Vol 40 No 6 (15 Mar 1982).

Amnon Yariv- Introduction to Optical Electronics
(Holt, Rinehart & Winston, Inc 1971).

- Quantum Electronics 2nd ed. (John Wiley & Sons,
1975).

- "Four Wave Nonlinear Optical Mixing as Real Time
Holography" Optics Communications Vol 25 No 1
(Apr 1978).

Amnon Yariv & Robert A Fisher- "Introduction"
in: [Fisher, 83].

B Ya Zel'dovich, N F Pilipetskii, V V Shkunov- "Experimental Investigation of Wave-Front Reversal Under Stimulated Scattering" in: [Fisher, 83].

- Principles of Phase Conjugation (Springer-Verlag, 1985).

CHAPTER 3

OPTICAL COMMUNICATION

Chapter Abstract - The inherently point-to-point nature of an optical data link fits well the interstellar communication problem. Binary digital encoding can transmit any desired message; link channel capacity emerges as the basic system performance criterion. Standard optical pulse-code-modulation specifications of 10^{-9} for bit error probability and 26 dB for signal-to-noise ratio are selected for the reference planetary laser design. A presumed matching receiver would collect incoming light with a large reflector, using heterodyne detection and tunable multichannel processing to find, characterize and track the Doppler-shifted signal, and discover its modulation rate. A fundamental link equation relates design parameters to the system channel capacity and hence its communication efficiency.

The basic difference between an optical communication channel and other, more familiar electromagnetic types is that the optical link's carrier frequency is several orders of magnitude higher. Radio frequencies range from order 10^3 to 10^8 Hz, and microwaves from order 10^9 to 10^{11} Hz, but optical frequencies, which include the infrared, visible and ultraviolet portions of the electromagnetic spectrum, range from order 10^{12} to 10^{18} Hz (Figure 3-1). We already see intuitively that an optical carrier should allow a greater information density to be impressed upon it, simply because its higher frequency provides more carrier cycles per unit time. It is this greater frequency bandwidth, with its promise of enhanced information capacity, which largely motivates the widespread present replacement of electrical cables by optical fibers for guided information transfer. Further advantages of guided optical systems are their inherently high security (since tapping is difficult to do and easy to detect) and ability to operate reliably in the presence of noise. Most of these reasons lead us to specify optical telemetry links within the planetary laser fleet developed in Part 2.

Clearly an interstellar link must be unguided, though, since no material connection is possible. In general an optical carrier will have [Gowan, 84]: lower generation efficiency than radio or microwave carriers, quantum-limited rather than thermal-noise dominated detection, a higher ratio of received-to-transmitted power, smaller system apertures, and a highly directed, rather than broadcast, nature. We suffer thereby the greater trouble of accurate aiming and tracking, but the nature of the interstellar communication problem should in fact benefit from a system "most directly suited to independent, point-to-point channels", because to first order star systems represent point targets in empty space.

For sending messages either to distant human colonies or to alien stellar civilizations, we can easily imagine using combinations of high-resolution color holographic and video images, acoustic images, and many, many symbolic images --- literary, mathematical, and numerical. All these varied types of information must be transmitted identically, as a time-varying electromagnetic function modulating the carrier laser beam in a detectable and decodable way. All information passing through the communication link is therefore strictly just a modulating pattern of electromagnetic data.

Digital Sampling

In common with most modern telecommunication systems, we choose to represent those data digitally. All practical signals are bounded-spectrum functions [Kaplan, 69]; that is, when mapped into frequency space by Fourier integrals, their component frequencies do not exceed the lowest and highest frequencies of some finite bandwidth Δf . Modern systems process signals with great versatility, and transmit them with improved reliability, by quantizing their continuously-varying (analog) signals, encoding them as patterns of discrete pulses (bits). Such translation is assured of losing no fidelity (in a "noiseless" system) if the original signal is sampled at a frequency f_s which is at least twice that of the highest-frequency component f_m it contains, where the range from 0 to f_m is the bandwidth Δf . Known as the Nyquist Sampling Theorem, this relation:

$$f_s \geq 2f_m = 2\Delta f \quad (3.1)$$

provides a lower bound on required sampling rate (a more practical criterion is $\sim 10f_m$, for real systems with noise).

We will assume for the reference planetary laser design the most common form of quantized, or digital, encoding: pulse code modulation (PCM), in which the only variable is signal pulse amplitude h , with all pulses having the same duration. The quantization gap is then the smallest difference in height Δh between possible pulses. The integral number of different pulse amplitudes allowed is called the base a of the system. The most common base of PCM systems is 2, yielding a binary encoding scheme. This is used, for example, in logic circuits where the current is either high or low, and in digital compact disc recording where a pulse is either there or it isn't. The number N of different quantum levels available to a binary coding scheme, and therefore the fineness with which it can represent a signal, then depends on the number m of pulses comprising a representational unit:

$$N = a^m \quad (3.2)$$

Thus a binary code, grouped in "bytes" of 8 bits each, can represent $2^8 = 256$ different symbols. While such a code works well for specific, prearranged symbolic communication (such as text) between knowing users, some desirable signals (such as a stereophonic, color, video representation of a symphony orchestra performance) require a larger palette than 256 discrete values if encoded simply. Now clearly the fineness N can be increased by increasing either a or m . Although binary is the least efficient coding base possible [Kaplan, 69], its common use derives from operational simplicity, since a logic gate can be made "high" or "low", for

instance. Information can be transmitted by switching a laser on or off, or deflecting its beam back and forth slightly so that a detector sees it only intermittently. Thus PCM systems generally increase the unit length m to achieve greater fineness.

Different types of information translate into different amounts of data. Assuming binary digital representation [Gowan, 84], an average 250 page book, for example, translates to about 3.5 Mb. Realtime digital voice channels require 64 kb/s. A typical audio compact disc requires 620 kb/s, and contains a total of roughly 2 Gb of information. Landsat data is downlinked at 100 Mb/s [NASA TB, 8702], and consists only of multispectral video from low earth orbit. Digitized color video transmission requires 142 Mb/s if encoded simply, with a 100 min film containing 850 Gb. Transmitting the yet more complex representations (such as moving color holographic images with stereo sound, of substantial duration) we might expect to produce in the future, to several target stars within a reasonable time, would require transmission rates much higher than current limited-resolution video.

Thus a large channel capacity is desirable in order to compress a lot of data into as short a transmission interval as possible. Even with a generous (by current standards) capacity of 10 Gb/s, we see from the above examples that a high-speed, compressing, digital binary channel would still take 0.2 s to send a Brahms symphony alone, or 85 s to send the information of only one Ridley Scott movie! A large channel capacity is also necessary to multiplex different "programs" (send them together on the same transmission line by interspersing their signals in time). The simplest relevant example is that the sender would typically duplex a CETI link, to transmit a repeating decoding tutorial continuously along with the main

messages. The channel capacity, then, measures the efficiency with which the system can transmit information with acceptable reliability, and is thus the real criterion by which we can judge the performance of an interstellar link.

Reconstitution and Errors

But "acceptable reliability" must be quantified. There exists a plethora of PCM codes with wonderful names like non-return-to-zero-mark, bi-phase-level, and delay modulation-space, each of which has different modulation rate requirements and total time-averaged transmitted power levels. We will focus for simplicity on return-to-zero signaling (RZ), in which energy is emitted as an impulse at some fixed time within a constant bit period τ [Gowan, 84]. Fixed optical energy ϵ_T is sent to designate a logical 1; sending no energy during τ designates a logical 0. In the ideal, noiseless case, the receiver samples the incoming waveform at the signal's digital sampling frequency f_s , in other words with the same bit period τ that the transmitter uses. If the received optical energy ϵ_R integrated over τ exceeds a programmed threshold level (which for simplicity we may assume to be half the rms peak level when a 1 is received) a 1 will be regenerated; otherwise a 0 will be regenerated. Applying a low-pass filter with cutoff frequency f_m to the regenerated signal then recovers the original information.

Real systems deviate from these idealized assumptions. Time variations when the bit period τ is not constant are called jitter. Amplitude variations when the received optical energy corresponding to a 1 is not constant are strictly

called noise. And the nonzero time duration of the optical impulse designating a 1 causes intersymbol interference, because energy properly belonging in one bit period overlaps adjacent periods. The problem of nonzero energy being received for a 0 is described by the extinction ratio r_e :

$$r_e \equiv \frac{\epsilon_R(0)}{\epsilon_R(1)} \quad (3.3)$$

which never vanishes in real systems.

All of these departures from the ideal case can cause a bit to be regenerated incorrectly. The probability of error PE in bit regeneration is defined simply using standard notation for a binary link:

$$PE = P(0|1)P(1) + P(1|0)P(0) \quad (3.4)$$

Since in an extended bit stream the probabilities of sending a 1 or a 0 in any given bit period are equal:

$$P(1) = P(0) = \frac{1}{2} \quad (3.5)$$

equation 3.4 can be simplified:

$$PE = \frac{1}{2} (P(0|1) + P(1|0)) \quad (3.6)$$

The degradations cited above all contribute to what we may call the total link noise, and constrain receiver properties, notably channel bandwidth. A wide enough receiver channel bandwidth can insure that a logical 1 signal enters the regenerator's decision circuit still as a short pulse compared to τ , but increased channel bandwidths let in more amplitude noise as well, and jitter worsens the probability of the decision circuit sampling an impulse signal off-peak. Both of these increase PE. On the other hand, a too-narrow receiver channel bandwidth may cut off some of the impulse response. From this we conclude that the optimum receiver channel bandwidth should match the linewidth of the transmission as closely as possible.

Pulse-to-pulse amplitude variations arise from nonconstancy in both transmitter and receiver, as well as interference from the transmission medium. Voltage (or current) variations due to thermal noise (Johnson noise) in the detector and the receiver electronics (presuming a non-optical-processing back end), are modeled by a Gaussian distribution about the mean value. Optical systems, because of their typically quantum-limited detection process, suffer mainly from signal-power-dependent shot noise, due to the randomness with which even a constant light flux generates carrier pairs in a photodiode, and modeled by a Poisson distribution. Total receiver noise consists of contributions from both kinds of noise, and due to their differing mathematical models, combining them predictively even when their proportions are known is not simple [Gowan, 84].

Interference in the transmission medium (interstellar space in our case) would come from scattering and absorption by interstellar matter. Attenuation by their transmission medium typically limits the range of optical communication systems;

the 25 pc distance we plan seems enormous, and silicate dust grains are known to absorb strongly at 10.6 μm [Hartmann, 83]. Astronomical absorption is greatest in the direction of the galactic plane due to its higher concentration of matter. For visible wavelengths along this worst direction, Allen [73] gives an extinction value of 1.9 mag/kpc (1.6 due to molecular clouds and 0.3 due to dust grains). The length unit of kiloparsec provides an immediate hint that absorption will not comprise a serious problem, since even our maximum design distance is 40 times smaller. Nonetheless, Appendix A3-1 proceeds with the calculation to derive a worst-case absorption of only a few percent. Infrared astronomers working with IRAS data assure us that our IR wavelength will suffer even less loss [Dwek, 86]. Since Allen's value is an average derived from observations over much vaster distances than those used here, it is clear that local interstellar extinction would vary greatly on a star-by-star basis among our target sample. In any case, such degradation need not be a major concern to this preliminary work, and we will subsequently ignore it.

A nonzero extinction ratio r_e derives from many sources, including intersymbol interference, any dark current present in the detector photodiode, and the presence of background sources which mimic the signal source. However, its major component typically comes from imperfect extinction at the signal source itself during transmission of a 0. Switchable semiconductor lasers biased near threshold commonly worsen r_e in fiberoptic systems, for instance. The modulation schemes we discuss in Chapter 7 for the planetary laser, because they merely deflect the constant output beam of a steady-state CW laser, can also never guarantee a zero extinction ratio.

The degree to which total link noise is tolerable determines the system's transmission reliability. Because of

the integrating nature of our perception, some types of information are much more noise-tolerant than others. That is, one bit error in a piece of text could change its entire meaning, whereas one bit error in a compact disc recording is often imperceptible. So although a transmitted message type represented by a lot of data will necessarily contain a comparable number of errors when received, if its representation is inherently redundant its message content will be error-tolerant. The only way a versatile system design can accommodate the range of error tolerances is, again, by considering all transmissions as just one data set, with just one worst-case error limit.

Depending on the system application, acceptable values for PE can vary from about 10^{-6} to about 10^{-15} , and Gowan [84] chooses 10^{-9} as "the normal PE requirement for a typical optical link." It might be argued that an interstellar link, with its atypically long return-verification delays, should specify lower error probabilities; on the other hand, Gowan arrives at his value presuming a guided fiber system in which external disturbances, not internal noise sources, are more likely to define performance by giving "rise to bursts of errors rather than a steady random distribution". Since the primary noise sources for a space link are intrinsic to the hardware, and can therefore be assumed in fact to follow that "steady random distribution", we relax predictions of interference and choose $PE = 10^{-9}$.

PE finds its way into channel capacity calculations via an associated quantity, the amplitude signal-to-noise ratio K . Also called the dynamic range, K is defined most usually as the maximum system signal amplitude divided by the rms noise amplitude. (Inherently dimensionless, it is often quoted in terms of a power ratio. Thus $20 \log(K)$ gives K in dB.) According to Gowan [84], for binary PCM systems K is

"relatively insensitive" to a precise choice for PE. Its exact derivation from the chosen value of PE depends on which mathematical model is used (the complementary error function $\text{erfc}(x)$ is common for systems dominated by thermal noise), but Gowan recommends $K = 20$ as a "very conservative" value for PCM system specification, one which includes a generous "unallocated system margin" to account for the uncertainty of accurately modeling probability distribution tails. We adopt $K = 20$ (26 dB) as a reasonable reference value for this study.

Presumed Receiver

Evaluating the performance of any interstellar transmitter we might design requires some knowledge of the link's receiver. If our mission is to communicate with a human-launched probe or a distant human colony, we get to design a "matched" system from the start. However, by definition SETI cannot know a receiver's characteristics, and CETI would, at least initially, not know them either. While there exist unanswerable arguments against an alien intelligence approximating our own to any substantial degree, we have no other sensible choice than to presume the type of receiver which we ourselves would use to detect the signal we intend to transmit. The bright side of that catch-22 is of course that it licenses us to presume a matched receiver, and one taking advantage of the same engineering abilities projected in Part 2 for the planetary laser itself.

Consequently the distant receiver for this study will use actively controlled, segmented mirrors to collect light from a large interception area of the source laser's far-field beam

pattern. The telescopic precedent for such a receiver exists already in the Keck telescope being built at Mauna Kea (with a 10 m segmented primary), and in a JPL baseline proposal for the space-based far-IR Large Deployable Reflector (LDR) (with its 20 m primary of 50 - 84 2 m hexagonal segments) [Mattingly, 86]. And a detailed systems precedent for optical performance of a truly large segmented structure is the planetary laser itself, in Part 2 of this work.

Since a receiving culture would not a priori know which of its nearest 1000 likely stars would be the most likely to transmit messages, it seems unreasonable to imagine it setting up 1000 1 km receivers, each trained on one star, to find out. But a few such facilities watching single stars for dedicated intervals is not at all preposterous. Indeed, the cost of such a search would be trivial compared to the cost of building the source transmitter, and quite affordable by a civilization capable of interplanetary engineering (Chapter 11). In any case, a really large receiver is only necessary after initial detection, to insure a signal-to-noise ratio sufficient to permit reliable detection of the signal content in a high-frequency-modulated beam. That is, just finding the beam is a quite simpler problem than receiving it "well" enough to extract its message bit stream. Discovering such a laser signal would undoubtedly encourage a searcher to establish a dedicated, better receiver. Both the point-to-point nature of a laser link and its potential for enormous data transfer rates thus argue for presuming a matched receiver --- laser CETI really only makes sense for target civilizations at least as technologically advanced as the sender. Consequently we presume a receiver diameter of the same order as that of the satellites producing the laser beam in the first place, a few km.

We next presume coherent heterodyne detection at the focus of that collector, because it is the most sophisticated technology we could apply. Such a device combines optically the collected source signal E_r with a well-characterized and much more powerful local oscillator signal E_l (produced in the IR case by a thermally stable, tunable mode-locked laser) as diagrammed in Figure 3-2. The requirement that such combination be accomplished coherently limits this technique currently to wavelengths longer than visible. Focusing the superposed waves on a HgCdTe photodiode cooled to LN₂ temperatures ($\leq 77K$) assures that the detector's noise contribution is minimized [Glenar, 81]. The output current I is proportional to the square of the input field superposition:

$$I \propto (E_r + E_l)^2 = E_r^2 + E_l^2 + 2(E_r \cdot E_l) \quad (3.7)$$

The cross-term is of interest because it reduces to:

$$E_r E_l \cos(\omega_l - \omega_r) t \quad (3.8)$$

which is easily separated electronically from the DC constant terms.

Thus a difference, "beat" frequency, called the Intermediate Frequency (IF), arises from the two input signals. Containing the original signal modulation, the IF is at a much lower, more manageable frequency than the carrier (typically but not necessarily GHz or lower for IR carriers), and so can be passed on to processing and recording datonics. Amplifier noise can be made insignificant compared to shot noise by increasing the power of the local oscillator; with sufficient

power (no more than a few mW is necessary, which can be produced by a hand-sized laser) a heterodyne system's sensitivity can approach the quantum limit of detection, depending on the presence of background noise radiation in the IF band. In fact though, Gowan [84] estimates that K for a receiver tuned to $10.6 \mu\text{m}$ and pointing directly at the sun would "be degraded by no more than 3 dB." Since a distant receiver aimed at our solar system would be pointing directly at our sun, we have good reason for choosing a generous "unallocated system margin" for K , as discussed earlier.

A degradation factor Δ_R which worsens K must be ascribed to the unknown heterodyne receiver. Conventionally taken as about 10, this allows for losses due to polarization-filtering (necessary for coherently heterodyning the local oscillator), chopping (necessary for calibrating against the noisy local oscillator), and optical train imperfections in the receiver apparatus. Δ_R can be reduced by special attention to good componentry and more complex equipment. For instance, employing a parallel system which passes both polarizations on to multiplexed detectors (which we must presume in any case to preclude chopping out half of the densely modulated signal) can regain 2 of the factor of 4 initially lost to those two sources [Mumma, 88]. Although the theoretical limit for heterodyne technology is about 3.1, we choose $\Delta_R = 5$, an eminently achievable value.

As the signal is received, it must be processed and recorded, so the receiver's channel bandwidth, the range in frequency space over which it looks for the signal, is critical to link performance. As noted earlier, a channel which covers too large a frequency range will degrade K by admitting a lot of noise along with the given signal, whereas a channel which is too narrow will amputate incoming signal power. Best performance occurs if the receiver channel width matches the

source linewidth. Chapter 7 shows how the coherence requirement of laser light combines with the huge size of a planetary laser to enforce an exceedingly narrow emission linewidth, of order just a few Hz. Nobody would build a receiver with such narrow channels, though, unless he knew or suspected specifically of planetary laser transmitters, as we shall now see.

The detected frequency of electromagnetic radiation depends on all relative motion between the transmitter and receiver parallel to their line of sight. The Doppler shift increases that frequency above the source laser's "rest" frequency when the transmitter and receiver move toward each other, and decreases it when they separate. Appendix A3-2 calculates the various Doppler frequency-shift contributions for our planetary laser if measured relative to nearby stars. Of the cyclical shifts, the greatest magnitude is about 3.3 GHz with a period of 225 d, due to Venus' orbital motion about the sun; the fastest is about 620 MHz with a period of 2 hr, due to the planetary resonator's orbital motion about Venus. This means that a receiver would have to track the 3.3 Hz-wide signal beam over a frequency range a billion times wider, which it could do without unacceptably degrading K by using simultaneously a billion adjacent 3.3 Hz channels.

Now multichannel signal analyzers with about 1000 times fewer channels are a current goal of NASA's SETI project, but there is no intrinsic reason why dedicated gigachannel receivers could not be made. In fact, the actual requirement would be much less daunting, for the periodic behavior of orbital Doppler shifts would allow a receiver to track them predictively after only a short characterization time, using a much smaller set of adjacent narrow channels. After the signal was found, the reconfigurable receiver would examine it using successively finer channels to analyze its actual width and

excursions. Studying the shifts themselves, of course, would provide the searcher with immediately decodable information about the transmitter's planetary system dynamics, and orientation with respect to the receiver. Expecting an alien civilization to anticipate the presence of an anomalously bright, Dirac delta spike somewhere within GHz of 28.5 THz (the frequency of a CO₂ laser) in the spectrum of a type G star seems like a reprise of the microwave "watering hole" dilemma cited in Chapter 1. Nonetheless, the notable major natural phenomenon of atmospheric CO₂ lasers makes this particular watering hole much less of a mirage, and we presume multi-narrow-channel reception.

Clearly the receiver must record the incoming signal with frequency f_r at least equal to $f_s = 2f_m$, or data in the bit stream will be lost and the message made meaningless. In the extreme case an alien searcher might find the signal but interpret it simply as a beacon, without noting its high modulation rate and the extensive information it contained! However, the signal consists in detail of more than a simple narrow peak at the carrier frequency f_c superimposed on background noise. Although its exact waveform will be complicated, the modulation produced by the reference CETI Transducer explained in Part 2 can be represented approximately by a sinusoid. Then the actual transmitted signal as represented in the frequency domain will contain spectral power at the sum and difference frequencies $f_c \pm f_m$ (where f_m is the modulation frequency), detectable as distinct peaks centered around f_c . A multichannel narrowband receiver which can locate and track the carrier peak could simultaneously detect these sidebands, whose separation would immediately reveal the proper f_r to use so as to match f_m and record all its encoded information.

The maximum possible modulation frequency f_m is practically constrained by datonic (electronic or photonic) signal-processing limitations. Although communications engineers are beginning to talk gleefully about all-optical amplification technology in the THz modulation range, and fully intend to leave systems based on electro-optically modulated lasers in the dust [Manneberg et al, 87], currently the fastest conventional electronic switch is an advanced modulation-doped field-effect transistor (MODFET) which can operate at a few hundred GHz [Weisburd, 86], and the fastest complete clock circuit is a GaAs chip at Hughes Research Laboratories which operates at 18 GHz [SN, 8701]. Thus a reasonably projected upper limit on electronically-constrained processing speed would be of order 100 GHz. Practical integration times required by state-of-the-art photodiodes for heterodyne applications seem limited to roughly a tenth of this, or around 10 GHz. Consequently Part 2 will presume 10 GHz as an upper limit for f_m , and hence f_s .

Appendix A3-3 assembles the key parameters of mesospheric IR emission, optical communication links, and heterodyne detectors into a fundamental equation which models the performance, measured by its channel capacity, of a planetary laser engineered for interstellar communication. Part 2 of this work does in fact engineer such a system, and evaluates its performance for CETI using this link equation. At the end of Part 3, this study concludes by returning to the link equation, establishing some theoretical limits and discussing their implications for the expansion of life throughout the Milky Way.

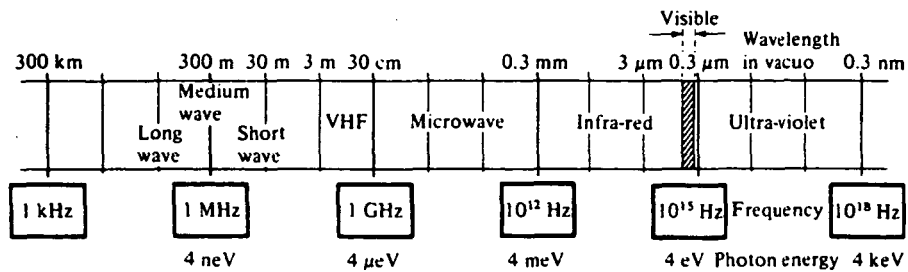


Figure 3-1 The electromagnetic spectrum. [Gowan, 84]

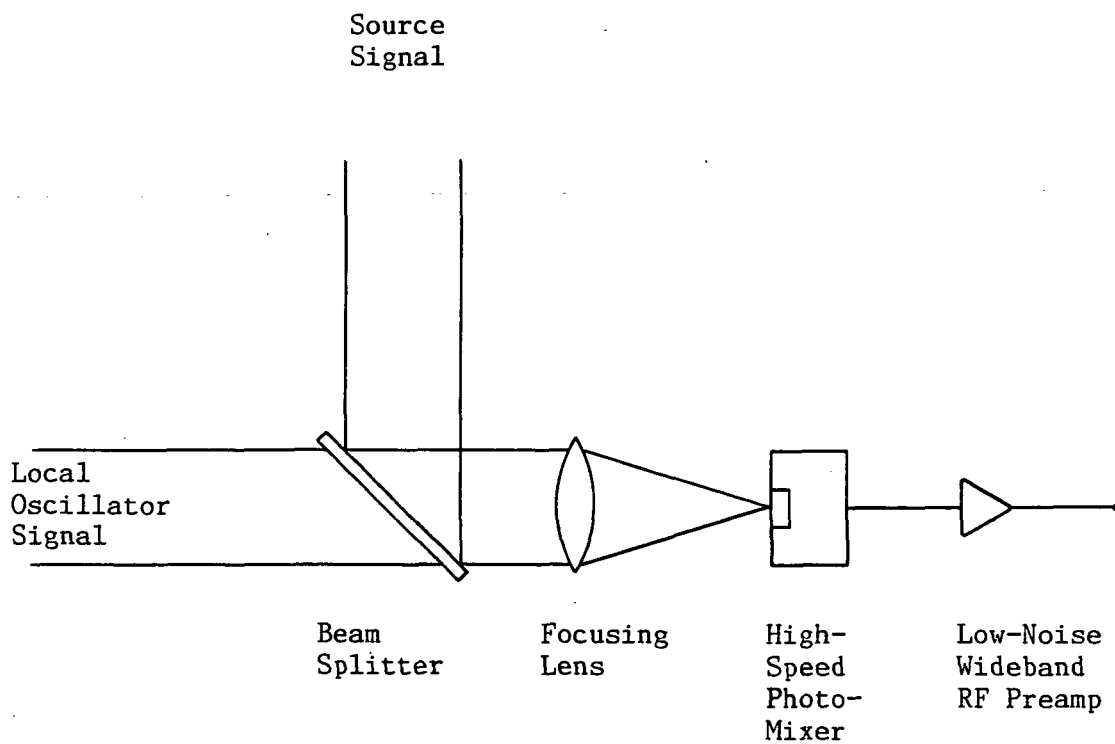


Figure 3-2 Laser heterodyne mixing. after Glenar [81]

Appendix A3-1 Interstellar absorption.

From Allen [73], visual magnitude is defined by the relation:

$$m_1 - m_2 = -2.5 \log \left(\frac{l_1}{l_2} \right) \quad (\text{A3-1.1})$$

where $m \equiv$ the visual magnitude of an object and $l \equiv$ its luminosity. If we now assume that the extinction of 1.9 mag/1000 pc applies homogeneously even to much shorter distances, we would expect an extinction of 0.0475 mag due to our maximum transmission distance of 25 pc. This in turn corresponds through equation A3-1.1 to a luminosity ratio of 1.04, meaning that 96 % of the light that started out actually made it through to the other end.

Appendix 3-2 Doppler shift for interstellar lasers.

First-Order Shift

There are three contributions, from the resonator's orbital motion about Venus, Venus' orbital motion about the sun, and the sun's motion relative to nearby target stars.

The resonator orbital speed is 6.52 km/s (Appendix A6-7), so the maximum resulting Doppler shift will have magnitude:

$$\begin{aligned}\Delta f &= f \frac{v_{\text{rel}}}{c} = \frac{c}{\lambda} \frac{v_{\text{rel}}}{c} = \frac{v_{\text{rel}}}{\lambda} \\ &= \frac{6520}{10.5(10^{-6})} = 620 \text{ MHz}\end{aligned}$$

The shift will vary approximately sinusoidally because only the in-line component of the circular orbit's velocity produces it. Thus if the line-of-sight is parallel to the orbit plane, the shift will reach its maximum positive value at the moment when the resonator coupler directly approaches the target, and its "maximum" negative value when the coupler recedes directly. The complete cycle takes one resonator orbit (2 hr) and is damped if the line-of-sight and orbit plane are not coplanar. In the limit that the target star lies on the line normal to the orbit plane, no shift results.

Venus orbits the sun at a nearly constant speed of 35 km/s [Wertz, 84]. The resulting Doppler shift will vary approximately sinusoidally as explained above, for the same reason (in fact, in exactly the same way since the resonator orbit and Venus' orbit are coplanar), reaching a maximum absolute value of:

$$\Delta f = \frac{35,000}{10.5(10^{-6})} = 3.3 \text{ GHz}$$

The complete cycle will have a period equal to the Venusian year, 225 d.

Finally, the sun's own motion relative to the stars differs from case to case, both because the line-of-sight component of the sun's inertial velocity varies with different target angles and because all the other stars are moving, too. The maximum relative velocity with respect to nearby stars [Wertz, 84] has magnitude 15.4 km/s, so:

$$\Delta f = \frac{15,400}{10.5(10^{-6})} = 1.5 \text{ GHz}$$

is the worst uncompensated Doppler shift that would result. We say uncompensated because, being constant and knowable from astronomical observations, this shift due to relative stellar motion can be removed by the receiver bias.

The contributions which necessarily affect the acceptance bandwidth of a reasonable receiver are therefore those cyclical ones due to orbital motion within the transmitter's star system.

Second-Order Shift

The second-order Doppler shift, which is sign-independent and proportional to v^2/c^2 [Richard, 87], is of course much smaller than those calculated above, being only 390 kHz due even to Venus' orbital motion. An interesting but small contribution is due to the random motion of the CO₂ molecules in Venus' mesospheric lasing medium. Their average speed can be found from:

$$v = \left(\frac{2kT}{m} \right)^{\frac{1}{2}} = \left(\frac{2(195)1.381(10^{-23})6.02(10^{23})1000}{44} \right)^{\frac{1}{2}} = 271 \text{ m/s}$$

where $k \equiv$ the Boltzmann constant, $T \equiv$ the gas kinetic temperature (Appendix A6-7), and $m \equiv$ the mass of a CO₂ molecule. The resulting

second-order Doppler frequency shift for the planetary laser amounts to 23 Hz. Maximum contributions due to the resonator's orbital motion and the sun's relative stellar motion are 13.5 kHz and 75 kHz respectively.

Appendix A3-3 The link equation.

We begin with the source laser. If the laser extracts energy from a relatively fixed portion of the atmosphere at a steady-state rate equal to the solar pumping rate (Chapter 5), then the available flux F can be represented most simply as:

$$F \approx \beta V \quad (\text{A3-3.1})$$

where $\beta \equiv$ the volume emission rate in photons/(m³s) along a tangential path through the mesospheric inversion layer, and $V \equiv$ the effective volume of the gain medium. Expanding V , we can rewrite equation A3-3.1:

$$F = \beta L_g \frac{\pi D_c^2}{4} \quad (\text{A3-3.2})$$

where $D_c \equiv$ the intracavity beam diameter and $L_g \equiv$ the effective gain length through the medium, a value which may be inflated beyond the geometrically-fixed single-pass gain length using a multi-pass resonator (Chapter 5).

Were the laser beam to exit the transmitter as a plane wave, its divergence in the far-field would be governed simply by diffraction spreading, and depend on wavelength, aperture diameter and transmission distance. Gaining fine aiming control over the beam by shaping the system reflector figures, however, makes far-field divergence a control variable. For communication missions out to at least the maximum range we consider in this study, 82 ly, the radius R_s of the source laser far-field diffraction pattern's central Airy disk can be chosen as a mission parameter (Chapter 7). Then that spot area is πR_s^2 , and the specific laser intensity it contains is:

$$P_s = \frac{\eta_s \eta_d \beta L_g \pi D_c^2}{\pi R_s^2 B_s 4} \quad (A3-3.3)$$

where $\eta_s \equiv$ the overall efficiency with which the laser's steady-state emission actually gets transmitted, $\eta_d = 0.84 \equiv$ the fraction of the transmitted beam's energy which is contained in the central Airy disk of its diffraction pattern, and $B_s \equiv$ the emission linewidth in Hz of the laser light. Equation A3-3.3 thus tells how spectrally bright the signal is when it gets to the target star system. A receiver of diameter D_r will intercept specific power:

$$P_r = \frac{\eta_s \eta_d \beta L_g D_c^2}{4 R_s^2 B_s} \frac{\pi D_r^2}{4} \quad (A3-3.4)$$

The link (power) signal-to-noise ratio SNR is defined for a heterodyne receiver [Mumma, 88] in terms of the received power P_r and a quantity called the noise equivalent flux NEF:

$$SNR = K^2 \equiv \frac{P_r}{NEF} = \frac{P_r}{\left(\frac{\Delta_R}{\sqrt{B_r \tau}} \right)} \quad (A3-3.5)$$

where $\Delta_R \equiv$ the receiver degradation factor, $B_r \equiv$ the receiver channel bandwidth, and $\tau \equiv$ the integration time over which received power is measured in that channel to detect one bit. For the optimal case where the receiver channel bandwidth B_r matches the source linewidth B_s , we can set them equal and substitute equation A3-3.4 into A3-3.5 to solve for the bit integration time required by a heterodyne receiver to detect, with signal-to-noise ratio K , the signal transmitted by a planetary laser:

$$\tau = \left(\frac{16 B_S R_S^2 K^2 \Delta_R}{\eta_S \eta_d \beta L_g D_C^2 D_R^2 \sqrt{B_R} \pi} \right)^2 \quad (\text{A3-3.6})$$

The number of bits per unit time which this integration time allows is $1/\tau$, so by introducing a factor of two to accommodate the Nyquist Sampling Theorem (equation 3.1), we can determine the useful channel capacity of the interstellar link, measured in bits of program transmitted per second (b/s):

$$B = \frac{1}{2\tau} = \left(\frac{\pi \eta_S \eta_d \beta L_g D_C^2 D_R^2}{16 \sqrt{2B_S} K^2 \Delta_R R_S^2} \right)^2 \quad (\text{A3-3.7})$$

References

C W Allen- [op cit Chapter 1].

C Baack (ed)- Optical Wideband Transmission Systems
(CRC, 1986).

Eli Dwek- personal communication, NASA Goddard Space Flight
Center, Laboratory for Extraterrestrial Physics (1986).

D A Glenar- Development of a Diode Laser Heterodyne
Spectrometer and Observations of SiO in Sunspots PhD
dissertation, Penn State (NASA TM-83902) 1981.

John Gowan- Optical Communication Systems
(Prentice Hall, 1984).

William K Hartmann- [op cit Chapter 2].

S A Kaplan- (ed) Extraterrestrial Civilizations - Problems of
Interstellar Communication translated from Russian:
Fiziko - Matematicheskio Literaturny, Moskva 1969. (Israel
Program for Scientific Translations, Jerusalem 1971).

G Manneberg, M Kull, S-E Lindquist- "An All-Optical Amplifier
for Terahertz AM Signals: A Proposal" Journal of
Lightwave Technology Vol LT-5 No 2 (Feb 1987).

Richard Mattingly- "Space Station Staging of the Large
Deployable Reflector (LDR): Interim Requirements Report"
JPL D-3182 (January 1986).

Michael J Mumma- personal communication, NASA Goddard Space
Flight Center, Laboratory for Extraterrestrial Physics
(1988).

Jean-Paul Richard- personal communication, University of
Maryland Department of Physics and Astronomy (1987).

Science News- "Speeding to a Gallium Arsenide Record" Vol 131
No 2 (10 Jan 1987).

NASA Tech Briefs- "Goddard's Massively Parallel Processor"
Vol 11 No 2 (Feb 1987).

S Weisburd- "The Fastest Transistors in the World" Science
News Vol 130 No 16 (18 Oct 1986).

James R Wertz (ed)- [op cit Chapter 6].

PART 2

PLANETARY LASERS

If you don't leave a trail of bread crumbs,
I can't tell what you were trying to do.

-- Michael D Griffin

CHAPTER 4

SYSTEM OVERVIEW

Chapter Abstract - The reference design developed and analyzed in the rest of Part 2 forms a usable laser from the natural Venusian atmospheric phenomenon. Six reflector satellites comprise the planetary resonator, working cooperatively to generate a 1 km diameter cavity field. A further set of three focusing and switching satellites removes an output beam from the rotating resonator, redirecting it to either of two modulating stations at Venus' collinear libration points. Each station modulates the beam, and uses one or both of its satellites to aim it properly at target stars. The baseline link capacity for unrequited CETI is in the range kb/s to Mb/s, a rate dependent on assumptions about the receiver's size, location and channel bandwidth.

The reference design of this work represents one possible planetary laser system. In this chapter, we tour the spacecraft fleet and its operation, following the laser beam from origin to transmission. For clarity, we present all systems as faits accomplis and largely devoid of critical subtleties, reserving the system details for appendices and the justifying tradeoff analyses of Chapters 5 through 9. Since a major premise of the design is its blend of extant, attainable and projectable technologies, readers seeking both a defense of what may appear as startling capabilities, and treatment of higher order subtleties, should refer to the appropriate sections of those subsequent chapters.

We consider a planetary laser built at Venus. Figure 4-1 first diagrams (to scale) the inner solar system, showing dimensions of the innermost planetary orbits, and their annual variations, compared to the size of the star which is our sun. The second diagram then enlarges Venus space (to scale) to show the locations of the L1 and L2 collinear Lagrange libration points of the sun - Venus gravitational system, and their annual variations, compared to the size of Venus itself. The final transmission stages of our laser system operate at L1 and L2; we will return to them later as we follow the beam on its way out of the system. The beam originates, however, in the Venusian mesosphere, so the satellites which spawn it orbit Venus.

When viewed in a rotating coordinate frame, the resonator cavity as formed by reflector satellites is pentagonal, with sides tangent at the 130 km altitude that maximizes total tangential gain in Venus' mesospheric CO₂ inversion layer (Figure 4-2). The 7641 km radius circular orbit circumscribing that pentagon lies in the plane of Venus' orbit around the sun, which makes it very nearly (within 30°)

equatorial with respect to Venus itself. Such orbital geometry allows a smoothly rippling and well-constrained solar-pumped laser at all times throughout the Venusian year.

The orbiting system performs two separate functions: forming the beam, and extracting some of it useably. Figure 4-3 diagrams to scale an optical path which can accomplish both functions, showing five vertex stations defining the rotating pentagonal cavity. The details of Figure 4-3 will become relevant as we cover each optical element in turn, but for now we use this figure to introduce nomenclature which recurs throughout all the chapters of Part 2. The planetary resonator refers to those six satellites labeled 1α , 1β , 2, 3, 4 and 5. The other three satellite systems of Station 1 (1δ , 1ϵ and 1γ) comprise the coupler switch. Nothing they do has any effect on the circulating cavity beam; rather they control and steer the output beam which 1β removes from the resonator, sending it in turn on to L1 and L2 Stations.

The Orbiting Resonator

Although at first the pentagonal path looks like a closed ring laser, it is not. Rather, the cavity is a linear laser oscillator which has been wrapped around the planet. Stations 1α and 1β , albeit adjacent and in fact structurally connected, are not optically linked, representing instead the two opposite ends of the closed cavity. By reflecting the laser light back along the linear path, they set up the oscillating geometry necessary for sustained laser amplification.

Stations 2, 3, 4 and 5, the basic vertex stations, are identical and all do exactly the same thing at their respective

corners of the pentagonal path: they keep the oscillating beam in resonant phase with itself and directed at the (orbital intercept points of) adjacent vertices through the proper mesospheric gas layer. Because they comprise the bulk of the planetary resonator, we use one to demonstrate the systems common to all the resonator satellites.

Figure 4-4 shows to scale an entire basic vertex station. The spacecraft, massing 95,000 MT and consuming 405 MW of electrical power, is a relatively thin (10 m) disk 1.7 km across, composed of four major systems: the reflector, the attitude controller, the power plant, and the propulsion plants. Some of the components of these are shown at larger scale in Figure 4-5. The reflector itself, which physically defines the 1 km diameter laser beam, is a 1700 x 1000 m elliptical array of 230,000 individually controlled hexagonal mirror segments, each 3 m across from vertex to vertex and 0.25 m thick. Made of honeycombed hot-isostatically-pressed (HIP) beryllium, they are optically smooth and monolithically rigid to IR tolerances, coated with gold on the front and anodized for thermal control on the back. The mirror segments do not touch anything, each being isolated but positioned from the back by three high-frequency electromagnetic (EM) "space bearing" actuators (Figure 4-6).

The powered, active parts of all these EM devices are in turn nodes of a redundant active truss (Figures 4-6 through 4-9), two bays deep. Each active truss member is a thin-walled tube of carbon/magnesium (C/Mg) composite, through which run power and intelligence lines. The stiff and strong C/Mg material has excellent thermal and electrical properties, and is anodized on the outside for proper radiative thermal exchange. The length, bending and local buckling behavior of these members is monitored by embedded fiberoptic dimensional and temperature sensors, and adjusted by segmented, tripartite,

high-frequency piezoelectric (PZ) ceramic films applied to the inner wall surface. Truss nodes are also active, setting the inter-member angles with PZ elements, and the low-frequency member length with interposed aluminum thermal actuators. The members and nodes latch together removably through standardized structural/utility connectors.

To allow lasing despite orbital motion and (mostly) thermal disturbances, each mirror surface must be continually maintained, with nm resolution, at an integral number of laser wavelengths' separation from the reference plane, a mathematical ideal established and updated several times each second for all the resonator craft by the fleet controller, the fleet's combined artificial intelligence. The EM isolation mounts, in concert with the active mirror support truss, perform this segment tuning, as well as the fine-pointing mirror tilt necessary for the satellites to track each other's non-Keplerian excursions. Predictive and adaptive control is coordinated by the optically interconnected fleet controller, a neural net capable of pattern feature extraction, memory learning, projective action and exquisite simultaneous motor control.

Modeled on the massively parallel hierarchy of organic brains, the controller works simultaneously on levels ranging from detailed to global. It controls mirror segments in groups of six surrounding a seventh. Each of those six segments also belongs to another group (Figure 4-6), coupling the overlapping groups' performance. Seven groups thus comprise a family; families join into clans, neighborhoods, regions, sectors and so forth. The active structure is arranged analogously, so that just several levels of control can organize each entire satellite and all satellites in the fleet. Despite its vast number of parts and trivial intrinsic stiffness, each spacecraft's mirror segments act as one phased IR reflector

with the support structure behind it a willfully rigid reaction ground for its tuning activity. Communicatively linked by wideband, dedicated intercraft optical lasers, the fleet of satellites acts as a single entity. The controller compensates for the intercraft lightspeed signal propagation delay with anticipatory motor behavior, based improvingly on its learned responses over thousands of identical orbits.

Myriad sensors provide the state data required for such unified behavior. Each mirror segment uses tiny accelerometers collocated with its EM mounts to monitor the accelerations they deliver. These and intersegment optical sensors get their modest power delivered photonically across the space gap from the bus. Both the front and back of each rigid mirror are useful for intra-family metrication, and feature integral retro-reflectors for that function. Sources and sensors are mounted throughout the support truss structure. Regional sensor heads get their vantage views atop light, actively stiff masts projecting well beyond both the front and rear segment surfaces (Figure 4-10). These masts are thin composite tubes, statically and dynamically shape-controlled by PZ surface layers according to strain data from fiberoptic sensors embedded in their ply layup. Sensors in the cavity beam are not endangered because its power intensity is less than 1 % of the total ambient sunlight intensity. Each region uses interferometers to measure its collective position relative to the cavity beam's resonant phase pattern, thus permitting the fleet controller to determine proper reference planes.

Accelerometers distributed throughout the bus structure, by providing independent inertial data, allow the fleet controller to verify its optical sensors over short times. In differential modes they also map in detail variations in the planetary gravitational field, enhancing the controller's knowledge of its orbital environment and its consequent ability

to compensate disturbances predictively. Optical limb and IR emission sensors enable the fleet intelligence to know where the mesospheric inversion layer really is, and fixed-head star trackers aiming out the starboard and port sides of the craft (toward the celestial poles) have a constant and clear view of distant stars for external attitude reference. Such inertial reference is necessary even though intercraft referencing is paramount for actually operating the laser.

The central reflector ellipse of a basic vertex station then, if positioned and oriented properly, and fed with electrical power, acts as an enormous, single, flat, diffraction-limited mirror, rigid to IR tolerances. That and the other mirror surfaces dispersed around the planet can then in turn act as one resonant cavity, stable to lasing tolerances. So that the cavity beam's 1 km diameter remains constant, the proper orientation for the basic vertex station reflectors is planet-facing, with their major axes parallel to the orbital velocity vector (Figure 4-3). That planet-oriented attitude is conditionally stabilized by the planetary gravity gradient; trim rotations to maintain that equilibrium attitude are performed by the craft's Annular Momentum-Control Devices (AMCDs), shown in sectional detail by Figure 4-11.

Each spacecraft has two AMCD rims, each an 850 m radius, thin Kevlar ring with embedded magnets and ferrite bands, positioned, constrained and spun electromagnetically by suspensor/drive stations spaced every 3 m along their circumference. The two rotate in opposite directions, so that their considerable individually stored angular momentum cancels. By tipping the rims electromagnetically, a nonzero control momentum develops, which necessarily turns the entire spacecraft until the rims are untipped. Combinations of rim tipping and speed changes allow full 3-axis maneuvers.

A redundant, active exoskeleton truss (of the same construction described earlier) encases the AMCD system to ensure its precise roundness and provide a repeatable ground plane for tipping reference. The suspensor/drive stations are themselves positioned adjustably by PZ mounts, and monitored with respect to the circular exoskeleton by optical sensors in the rim chase. The chase, containing all AMCD equipment as well as power and intelligence utilities, is shielded from debris impact by removable bumper panels of layered reactive metal foam (RMF). Mounting fittings are typically titanium. The AMCD thus comprises a structurally distinct subsystem from the active mirror plane, capable of performing its stabilizing function when fed with electrical power. Its control intelligence is of course enmeshed with the mirror control intelligence at high levels, so that the systems act cooperatively.

The power for both mirror and AMCD control, and the structural connection between them, both derive from a web of power plant modules bridging the crescent gaps between the reflector ellipse and its circumscribing AMCD bus (Figure 4-12). A controlled, refractory metal nuclear reactor core, fissioning uranium nitride fuel, occupies the center of each power module, shielded by tungsten and lithium hydride and held in place by carbon/carbon (C/C) structural vanes containing parallel lithium heat pipes for its primary thermal transport (Figure 4-13). These give up their heat to a C/C composite tube 10 m in diameter, which is also the temperature-stabilized power plant structural armature, surfaced conductively to provide electrical grounding for thermoelectric (TE) converters. Silver radiator panels wrap the hot tube, stood off from the ground strips by the semiconductors which are the heart of TE conversion.

The power plant tubes are spaced far enough apart to allow practically unity radiative view factor for their almost 365,000 m² of high-temperature radiators. Power collection trunks from separate modules meet through regulator junctions at their structurally latched ends. The power plant structure is quasi-active; although embedded fiberoptic strain and optical displacement sensors monitor their relative motions, the tubes themselves operate much too hot for PZ actuators to survive. Thus all structural control occurs in the active C/Mg truss members which, thermally separated from the converter tubes by titanium attachments, join the ends of the reactor module assemblies to the AMCD and mirror bus structures (Figure 4-14).

This now powered and stabilized, optically precise reflector satellite is kept in its proper position along the resonator orbit by four xenon-ion engine plants spaced cardinally around its circumference. Each plant uses four arrays of 88 high-performance 50 cm engines (Figure 4-15) to develop sufficient thrust for countering solar gravitational forces and solar pressure, for other routine station-keeping, and for desaturating the AMCDs when necessary. Enough xenon is tanked in easily replaceable composite pressure bottles to last for about a decade, of the same order as the typical reactor module lifespan.

Appendix A4-1 develops consistent power, mass and inertial property estimates for Stations 2, 3, 4 and 5 based on a more detailed examination of the ships' subsystems and their interrelated functions. Appendix A4-2 then analyzes the attitude control and propulsive authority, and mirror fine-pointing performance, available from the subsystems outlined by A4-1.

The cavity end reflectors, 1α and 1β , differ from basic vertex stations primarily only in detailed configuration. First, they are smaller because they intercept the intracavity laser beam less glancingly. 1α , in fact, is essentially normal to the beam; hence its reflector need be only a 1000 m disk. Its mass, scaling in proportion to its mirrored area, is therefore only about 65 % that of the larger craft. Similarly, 1β , whose reflector is slightly elliptical (1000 x 1100 m) has about 70 % of the large mass value. Without detailed subsystem design, we allow interstitial room for their reactor plants as before, assuming a reference AMCD radius (and thus total satellite radius) of 650 m. The second difference for 1α and 1β is that they fly connected by a quasi-active (including passive members) truss structure, which maintains simply their non-principal-axis orientations against the constant Venusian gravity gradient torque, and resists the substantial gravitational attraction they feel for each other.

Extracting Output

Being also the laser's output coupler, 1β has a third difference as well. Its gold mirror surfaces are ruled by a diffraction grating which both ensures that the planetary CO_2 laser oscillates on a single spectral line (the P12, at $10.513 \mu\text{m}$ wavelength) and scatters 2 % of that circulating cavity power out of the resonator and in fact out of the orbital plane altogether. This 180 kW output beam can then be manipulated by optics which, albeit diffraction-limited in their IR accuracy, need not work in joint phase across thousands of km of space to maintain temporal coherence as did the resonator satellites. Coherence length ceases to matter for the beam extracted from the resonant cavity. Its

light now constitutes just a bright source of extreme spectral purity, perfect for deep-space communication.

The 1 γ satellite captures that light, concentrating and redirecting it on to the switching satellites. With about the same area as 1 β , the deeper 1 γ has a segmented reflector surface configured as a concave off-axis paraboloid, to converge the 1 km collimated output beam, thus making it manageable later by smaller optics. Its non-coplanar orbital position (Figure 4-16) is maintained by balancing counter-mass distributed across the orbit plane from it (not shown in that optical path diagram), and connected to it by quasi-active truss structure. The combined assembly rides slightly below the resonator orbit, so its mass center is brought up to the orbit by tethering more counter-mass at a higher altitude. Controlling the attach point of that tether force helps stabilize 1 γ 's non-principal-axis orientation, as does the gravity-resisting quasi-active bracing structure separating it from 1 β . The counter-masses referred to here consist of the spares and stores warehouses, robotic repair shops, and maintenance robot docks for the entire fleet.

The enormous size of the large Station 1 satellites, causing both their extensive geometrical distribution about the mathematical resonator orbit altitude and their great mass, prevents any portion of Station 1 from being in strict free-fall. However, the μg accelerations which do perfuse those satellites are easily accommodated by their structures, actuators and isolation mounts.

Station 1 is completed by two small satellites, 1 δ and 1 ϵ (Figure 4-17), a tethered pair stabilized by the gravity gradient and threading the structured gap between 1 α and 1 β (Figure 4-3). Together they comprise a recollimating directional switch to collect the narrowing beam from 1 γ , undo

most of its convergence, and keep the resulting 10 m diameter intermediate beam targeted alternately on the two modulator stations at the distant libration points. Tethering them together enables $l\delta$ and $l\epsilon$ to fly, passively, below and above the resonator orbit respectively and thus remain in formation with the rest of Station 1 despite their optically-required non-Keplerian positions.

The 50 MT $l\delta$ (Figure 4-18) contains a 550 kW out-of-core thermionic nuclear power plant (which supplies both craft), along with its roughly 500 m² of C/C high-temperature radiator. A redundant active truss of standard fleet construction connects this plant to the "payload" end of the bus, a 15 m convex off-axis paraboloidal reflector comprised of 1 m hexagonal, gold-surfaced beryllium mirror segments (Figure 4-19). These are supported and controlled by an actuated structure similar to that found in the resonator reflectors. Surrounding the reflector dome is a collar of actively stiff masts supporting optical figure sensors to monitor the reflector's shape, and overspill sensors (photodiodes tuned to 10.5 μ m radiation) which, through dedicated intercraft optical datalinks, enable $l\gamma$ to keep the beam focused on $l\delta$ at all times by microtilting its own mirror segments. 15 m dual AMCDs counter-rotate in a protective RMF chase around the reflector periphery.

Surrounding and occupying the bus mass center (CM) is the tether bearing, a spherical device nested in the active bus structure. Conductive power tethers penetrate its core, gripped by wheels which enable $l\delta$ to crawl along them while maintaining electrical contact. The power cables feeding the tethers and mechanism are flexible; EM space bearings support the entire assembly, isolating its vibration from the rest of the craft and permitting large relative rotations. These rotations accommodate the non-principal-axis orientation of $l\delta$,

necessary for beam targeting and produced by offsetting the combined tether axis from the craft CM. PZ actuators effect this offset by moving the bus' portions of the space bearings. 16 20 cm xenon-ion station-keeping engines with their associated tanks are located around the AMCD chase; the other 8 required for redundant 6 DOF propulsion are mounted on the bus structure which rings the tether bearing.

A 15 MT inert counter mass hangs 200 m below $l\delta$, at the bottom of the 3 MT tether system. The short tethers themselves are multiply redundant and spaced apart to avoid satellite separation in the event of catastrophic debris impact. Each is braided of Kevlar microstrands impregnated with copper for power conduction and surfaced with nickel to prevent gradual tarnishing in Venus' tenuous atomic exosphere.

1800 m above $l\delta$ and stabilizing the top of the tether system (Figure 4-20), the 55 MT $l\epsilon$ craft has the job of directing the intermediate beam away from Venus and on to the modulator stations. Its bus configuration is dominated by the circular redundant active exoskeleton which houses its 100 m diameter AMCDs. Their chase construction is just like that of the resonator satellites, but 50 % smaller in section and of course much smaller in diameter. 157 suspensor/drive stations are located at 2 m intervals around its circumference. The ring bus supports four quad 20 cm xenon-ion engine outriggers spaced equally around its periphery, allowing full 6 DOF propulsive maneuvers. The tether attachment mechanism at the "bottom" of the ring features 2 DOF actuators which adjust the CM offset to tame external torques normal to the tether axis.

The bus also supports as payload two identical, diametrically located, mirror assemblies (Figure 4-21). Each has a 10 x 15 m elliptical, monolithic honeycombed beryllium mirror surfaced with gold, mounted for fine-pointing control

from the back by three noncontacting EM space bearings. IR overspill sensor collars provide the information δ and γ need to keep the intermediate beam centered properly. Inter-craft telemetry telescopes and lasers are mounted directly on the bus, but all aiming and star-tracking sensors are mounted around the mirrors' rims so as to be collocated with them; small dedicated radioisotope thermoelectric generators (RTGs) power them since they are physically isolated from the spacecraft bus. The RMF and titanium mounting armatures in turn mount with rotating linear PZ motors on fixtures attached to the bus, allowing the mirrors to pivot about axes parallel to the bus ring axis. With this geometry, ϵ can retarget the intermediate beam to any point in the orbit plane.

The L1 and L2 points are of course in the orbit plane, since they lie in the plane of Venus's orbit about the sun. As Station 1 crosses the planetary terminator onto the dayside, δ tilts slightly to retarget the beam from the trailing to the leading mirror on ϵ . Simultaneously, ϵ rocks in-plane slightly, so that after they pivot its rim mirrors together can retarget the beam to L1. This entire switching maneuver takes 20 s. The system tracks L1 uninterrupted for a quarter orbit (about half an hour), when ϵ 's mirrors reverse primary and secondary roles. Another uninterrupted half hour later, Station 1 crosses the planetary terminator onto darkside, and the process repeats, this time with L2 as the penultimate target. The tethered pair thus keeps the intermediate beam virtually continuously trained on the twin modulators, thereby sustaining a usable communication carrier at all times throughout the resonator orbits and the Venusian year.

Appendix A4-3 details consistent power, mass and inertial property estimates for the δ - ϵ tethered pair, while Appendix A4-4 analyzes the resulting critical performance

available from their tether, attitude control and propulsive systems.

Modulating and Transmitting

L1 Station and L2 Station co-orbit the sun with the Venus system. By occupying the near-planet collinear Lagrange libration positions, these twin stations remain respectively subsolar and antisolar, so that one or the other is always visible to the 1ϵ switch. They thus alternate, twice per resonator orbit as described above, in capturing one fifth of the intermediate beam's diffracted energy, impressing it with the mission signal, and retargeting it toward stellar destinations. Each station consists of two craft, a Transducer and a Ring.

The Transducer receives the incoming beam from 1ϵ , removes its periodic angular pointing oscillations caused by the switch's orbital motion about Venus, smoothes its envelope, modulates it, and aims it either directly at targets in half the sky or at the Ring (which then aims it at the other half of the sky). The Transducer for the reference CETI mission, as shown in Figure 4-22, is a 15 MT disk 15 m in diameter and one tenth that thick, comprised of two independent portions. The forward section (Figure 4-23) is dominated by a $100\ \mu\text{m}$ thick modulator membrane of amorphous evaporated beryllium surfaced with evaporated gold and explosion-bonded (simply supported) to the continuous lip of a circular beryllium frame. Membrane shape is monitored by fiberoptic strain sensors bonded to its back face. Stiff compared to the membrane, the frame's shape, monitored also by embedded fiberoptic sensors, is

actively maintained by piezoelectric strain actuators bonded to its surface.

This modulator works by rapidly yet subtly bending the reflecting membrane from behind to effect amplitude modulation. As the beam is dispersed beyond its programmed defocus bias by the doming membrane, the laser signal value seen in the far field changes from logical one to logical zero. A centrally located electromagnetic transducer distorts the membrane for this purpose by reacting against the frame. Its bias position and oscillation envelope are determined according to mission specifications by the fleet controller, which also produces the actual modulation signal according to the transmission content, downloaded ultimately through dedicated optical laser links from separate, obviously manned facilities.

A multifunction 7.5 m collar of IR overspill sensors halos the circular frame. First, it monitors the laser beam amplitude to avoid impressing data on an off-nominal carrier; transmission resumes after the transient retargeting intervals only when all system components are properly aimed. Second, by measuring the diffracted intermediate beam's central Airy spot size and offset, the sensor network enables the Transducer to provide continual fine-pointing feedback to the switching satellite group back at Venus. This vital information, transmitted by optical laser link, is outdated by less than 2 sec due to the light propagation delay between Venus and the libration points. Finally, applying the actual Airy spot offset as a corrective bias to the Transducer's own targeting removes errors accumulated earlier in the optical train, assuring stellar targeting as accurate as the controller's onboard systems permit.

Small optical lasers and receiving telescopes for interstation telemetry, as well as aiming star trackers, are

mounted directly around the frame rim. The ship's complement includes both fixed-head and movable trackers, redundantly useful no matter what the beam exit angle is, since the all-sky target star catalog is maintained in processed memory (the mirror does not "face" its target). The brain centers of the fleet controller which are dedicated to working with the mission signal are located on the rear of the frame.

The aft portion of the Transducer, comprising the spacecraft bus, is connected physically to the forward section only by centrally located (momentless) flexible power cables for maximum dynamic isolation. Optical links transmit intracraft telemetry across the gap, whose three DOF are controlled by high-frequency EM space bearings. Figure 4-24 shows a fourfold radial bus symmetry. At the bus core (Figure 4-25), arranged to occupy the spacecraft mass center, are inertial sensors. Overlapping information from these instruments, the strain sensors, the EM isolation mounts, and the forward star trackers enables the controller to derive at all times the Transducer's high-frequency global and differential attitude state.

Immediately aft of the inertial package, a compact in-core thermionic nuclear reactor, fissioning uranium dioxide, cooled by alkali metal, and heavily shielded by tungsten and lithium hydride, provides a steady 30 kW. Control and power conditioning systems double as ballast mass just forward of the inertial core. The efficiency of the in-core thermionic converter, and its high rejection temperature, allow a small (8 m^2) C/C heat pipe radiator, configured as a shallow aft-facing dome. To permit maintenance changeout the entire reactor/radiator assembly can be withdrawn from the bus and replaced quickly.

A lightweighted RMF bus armature holds the reactor and extends outward from it, actively shape-maintained by segmented surface PZ films. Clustered around the reactor hub are redundantly interconnected, composite high-pressure xenon propellant bottles, easily replaceable and holding a ten year supply. Both the propellant and power lines follow the armature spokes out to four ion engine modules, each with four 20 cm station-keeping and angular-momentum desaturating ion engines. Two re-entrant channels at the armature rim house the dual AMCD rims with their 24 PZ-positioned suspensor/drive stations; the inboard side of each channel is closed by a lightweight composite debris bumper, removable for system maintenance.

The AMCDs turn the Transducer back and forth as it tracks Station 1 across half of the 15 mrad angle from one terminator crossing (acquisition) to the other (loss of signal), cancelling that motion's contribution to beam targeting. The EM modulator mounts always fine-tune the reflector's attitude at high frequency to allow nrad interstellar pointing accuracy. Appendix A4-5 outlines power, mass and inertial property estimates for the reference systems design of the baseline CETI Transducer, and Appendix A4-6 investigates the critical attitude and propulsive authority available from those systems.

The Transducer's diaphragm modulator mirror is sized to reflect a minimum beam diameter of 10 m at 45° incidence angle. Thus with the exception of target stars hidden for a few days (by the sun for L2 and by Venus for L1) every year, each Transducer can aim its signal anywhere in the 2π sr half of the celestial sphere centered on its view of Venus (the beam source). The complementary Ring (Figure 4-26), when in use, reflects the signal to targets behind its Transducer, allowing

full-sky continual coverage as the libration stations take turns in every resonator orbit.

The 75 MT Ring straddles the Venusian beam on its way to the Transducer. Extremely similar in configuration to 1ϵ , the Ring's AMCD armature is identical, but the diametrically located payloads are in this case a single secondary mirror, and the spacecraft power plant. As the Ring pivots around the incoming beam, its simply pivoted mirror can redirect the now-modulated laser to any point in the 2π sr hemisphere behind its Transducer (except those blocked by the sun from $L1$). That 12 m diameter mirror (Figure 4-27) is a monolithically rigid, honeycombed beryllium plane blank surfaced with gold. A small sensor collar around, and inertial sensors contained within, the mirror perform the by-now familiar job of gathering collocated alignment and collimation data. Those sensors, powered by a dedicated RTG, beam their data optoelectronically across the mirror's EM isolation mount gap to the bus. Compromised by no mechanical contact whatsoever, the space bearing actuators fine-tune the secondary mirror's three DOF for interstellar aiming in concert with the Transducer's mirror control.

The bus side of the EM mount, a titanium and RMF armature, can in turn pivot through a 45° stroke along a circular track centered on the mirror face and oriented parallel to the Ring radius. This motion permits the laser to exit at Ring "elevation" angles between 0° (parallel to the Ring) and 90° (normal to the Ring); combined with 2π Ring "azimuthal" bus rotation, it directly allows targeting the entire 2π sr hemisphere behind the Transducer. The track is a large piezoelectric linear motor, capable of reliable, backlash-free actuation at resolutions well within the EM mount's ability to perform fine-tuning. The fleet controller's mission-targeting brain center completes the active payload, and ballast mass

comprises a passive payload to balance the power plant mass across the Ring.

The power plant (Figure 4-28) is a shielded 300 kWe out-of-core thermionic nuclear reactor fissioning uranium nitride and cooled by liquid lithium. Surrounding the converters is a finned, radial C/C composite heat-pipe radiator, constrained to reject heat primarily in directions seen neither by the bus, the secondary mirror nor the Transducer. The reactor control mechanism caps the assembly for maintenance access, and the power conditioning system doubles as a bus interface platform. Although the power plant is inherently more massive than the secondary mirror assembly, their mass moments must balance to keep the spacecraft centroid on its symmetry axis; this is done most simply by ballasting the lighter of the two.

An actively stiff laser beam sensor web, similar to the Transducer's halo, spans the Ring's enclosed area but leaves a central 42 m diameter clear opening. Once the Transducer has acquired the Venusian beam, information from this sensor web enables the Ring to remain centered in the beam's diffraction pattern, presenting a steady target for the Transducer without obstructing the latter's view of Venus. Telescopic sun sensors and inertial sensors distributed around the Ring provide attitude state data. Station-keeping thrust is generated by 16 redundant 20 cm xenon ion engines mounted to the bus exoskeleton at four stations, each with its own 12 yr xenon supply. Neither the Ring's operation nor its extremely benign field environment should introduce accumulating secular attitude torques, but the ion engines are capable of desaturating out-of-plane torques. The tremendous rotational energy stored in the AMCD rims precludes a need for spin desaturation in the decades anticipated between overhauls,

since the fleet controller programs retargeting spin maneuvers to cancel over long times.

The Ring systems, with their power requirements, mass and inertial properties, are outlined in detail in Appendix A4-7. The consequent attitude and propulsive station-keeping performance is analyzed in Appendix A4-8.

Only one modulator station at a time uses its Ring. For those times when a Transducer targets the receiving star directly, its Ring must in general be stored out of the way of the transmitted beam. That means the Transducer and Ring must exchange places with respect to Venus. The ion engines on the lighter Transducer then move them together, until the Transducer passes through the open gap in the Ring's center (with 6 m clearance all around). Complementary engines then decelerate it for station-keeping, now "in front of" the Ring. This maneuver takes about an hour. Together then, the twin pair-formation L1 and L2 Stations can send the Venusian laser beam, modulated with a desired signal, to practically any star in the sky, with a pointing accuracy dominated by the desired target spot size, almost continuously for periods between hours and years.

Assembly and Maintenance

Completed parts for the modular fleet craft arrive by interorbital electric tug from their distant factories near 1 AU, and are stockpiled close to their destinations in Venusian orbit and at L1 and L2.

The Transducers arrive in only a few pieces because of their small size, but accompanied by several duplicate spares which are depoted in halo orbits around the libration points along with spare parts for the Rings. Similarly, 1δ arrives only slightly dismantled at Station 1. Its spare parts, and those for the rest of the orbital stations, are depoted as part of 1γ's ballast masses. After 1ε is built nearby, the conductive tethers are threaded through 1δ from below and attached to 1ε. Then the inert counter-mass is fixed at the bottom of the system, and 1δ and 1ε slowly separate as 1δ crawls downward. The ships' propulsive systems keep the combined mass center properly stationed at the resonator orbit altitude until deployment is complete.

Because the craft are designed to be easily serviceable by robots, they are generally easy to assemble in the first place. A few special requirements deserve attention, however. Being monolithic, the Kevlar AMCD-rims are really the figurative core of each spacecraft, for it is around them that everything else must be assembled. A proper sequence for the large satellites starts by building the outboard, "top" and "bottom" sides of the AMCD exoskeleton. This structure requires active control during the entire operation; its members, after all, are sized presuming powered control, and the dynamic loading during assembly probably exceeds any they will experience later. Power at this stage comes from soft-docked, temporary construction reactors added on as needed. Precise control is unnecessary; the structure so far needs only enough authority to keep from tearing itself apart. Global attitude and propulsion should still be irrelevant at this time, too; attached control units can provide any temporary authority necessary. The rim hoops can then be installed, and the suspensor/drive stations mounted into the chase framework. Because of the enormous energy they eventually store rotationally, powering up the rims will take days, commensurate

with the limited power rating of the construction reactors. As the rims expand radially upon speeding up, their nominal performance can be inspected carefully. Finally the inboard chase panels can be installed, and the exoskeleton completed.

With propulsion plants and xenon tanks added, such an AMCD bus can now "take care of itself" positionally. If destined for a Ring, once its own power plant and payload mirror assembly are installed, and full sensor complement and nervous system are rigged, its construction is complete and control is turned over to its initial program. If destined for 1ϵ , two mirror assemblies are installed and the power tethers attached before the spacecraft comes on line; then the $1\delta - 1\epsilon$ pair deploys to proper station. If the AMCD bus is for one of the really large craft, though, it is far from finished.

The construction reactors for those large craft are of course several of its individual TE modules. Once the AMCD is complete and the ion engine plants attached, other power modules are mounted around the inboard edge and brought on line. Then the soft-docked construction power plants can be detached and moved around to extend the power web further toward the bus center. The mirror support truss is assembled into groups and families which then, when mounted along the elliptical inner edge of the power web, start the reflector plane growing toward the spacecraft center. Attaching the mirror segments themselves, which occurs after their supporting structure is in place, powered and checked out, is simple since they do not physically touch their mountings. Surface figure sensors are mounted as each neighborhood is assembled, and of course every component begins its active life as part of the fleet intelligence as soon as it is connected. By the time an entire craft is completed, its controller already knows in great detail how the ship behaves.

By that time also, components will have begun failing as well. The job of replacing and repairing damaged or failed pieces of the fleet, restoring degraded materials, and changing out depleted power plants, is performed throughout its mission life by an extensive subfleet of variously specialized itinerant robots. Some are strong, others dextrous; all are free-flying spacecraft under control of the fleet intelligence. Although they themselves rove endlessly about the fleet craft, the repair-shop hangar to which they bring recyclable components is the tethered counter-mass for 1γ . Control and power hardware used during the fleet construction gets reconditioned here into elements of the maintenance subfleet.

Once all craft in the fleet are fully assembled and in control of their own performance, they establish the laser datalinks among them which awaken the true fleet intelligence. Immediately the controller begins refining its initial program with the performance and environmental data pouring in; noting patterns and testing responses, evolving and streamlining procedures for operating the planetary laser under increasingly familiar conditions. After a sufficient learning interval (minutes? days?), the controller can bring the orbiting craft to reference-plane performance as a resonant cavity, thus striking the laser. Finally, with all coupler, switch and modulator craft in alignment, the fleet is ready to download a targeting program and signal bit stream, and send its bright, spectrally narrow, modulated beam toward the stars.

Link Performance

Having a reference design for the planetary laser enables us with some authority to substitute reasonable values in the parametric link equation (A3-3.7) and thus evaluate how well the system performs the CETI mission. The tangential subsolar volume emission rate β is taken as $2(10^{13})$ ph/(m³s) [Deming & Mumma, 83]. The effective gain length is 600 km, which is 50 % greater than the single-pass interaction length to accommodate the gain enhancement introduced by a pentagonal resonator (Chapter 5). The cavity diameter D_c is 1000 m. In Chapter 3 we chose a matched receiver diameter (1000 m), a signal-to-noise amplitude ratio K of 20, and a receiver degradation factor Δ_R of 5. η_s is 0.21, the fraction of the diffracted intermediate beam's energy (a fourth of its central Airy spot) which the Transducers intercept and can retarget. η_d is just the energy fraction left in the central Airy spot of the far-field pattern, 0.84. The FORTRAN program LINKPERF listed in Appendix A4-9 embodies the link equation using these values. Although the laser's emission linewidth is 3.3 Hz (Chapter 7), LINKPERF retains the receiver channel bandwidth as an independent variable in order to reveal the effect of detecting the signal with an unmatched receiver. The channel bandwidth calculation is then parametrized according to the target spot size chosen with a shape bias of the Transducer's modulator membrane.

Figure 4-29 graphs the result. For a receiver whose channel bandwidth matches the emission linewidth, a Mars-orbit sized target spot is sufficiently small to yield Mb/s data transfer rates with our reference error probability (10^{-9}). Spreading the beam over a larger spot size clearly reduces the attainable data transfer rate, as does increasing the receiver channel bandwidth beyond the matched condition. Conversely, targeting an orbit around the star comparable to Venus' or

Mercury's enables transfer rates to a matched receiver which obviously would exceed the reference CETI Transducer's electromechanical modulation speed.

The performance model indicates quantitatively the point-to-point advantages of a laser system for interstellar communication. The data transfer rate can be made quite large if, first of all, the target location is predicted accurately, accomplished by aiming the laser to a small spot centered on the target star. Thus a strategy for successful SETI presuming planetary lasers must include establishing a large receiver in one's inner solar system, because that is where an optimistically high-frequency-modulated signal would be aimed. But second, the data transfer rate can only be made large for a matched receiver. Since a civilization aiming unrequited informational signals at other stars would be presuming an adequate reception facility, the successful SETI strategy mentioned above must also include a reconfigurable multichannel receiver capable ultimately of matching the exquisitely narrow spectral line a planetary resonator is constrained to produce.

The target-size dependence in particular would benefit greatly from return- or fore-knowledge of a receiver's location. Two-way CETI thus leads naturally to enhanced data transfer rates for a laser system, as does the mission of communicating with human colonies, where a truly matched system could take full advantage of the inherently directed nature of a laser link. In Chapter 12 we extend the present result to show just how well such an intra-species interstellar link could perform. But first we must justify in some detail the reference laser system, simultaneously laying the groundwork for that enhanced mission.

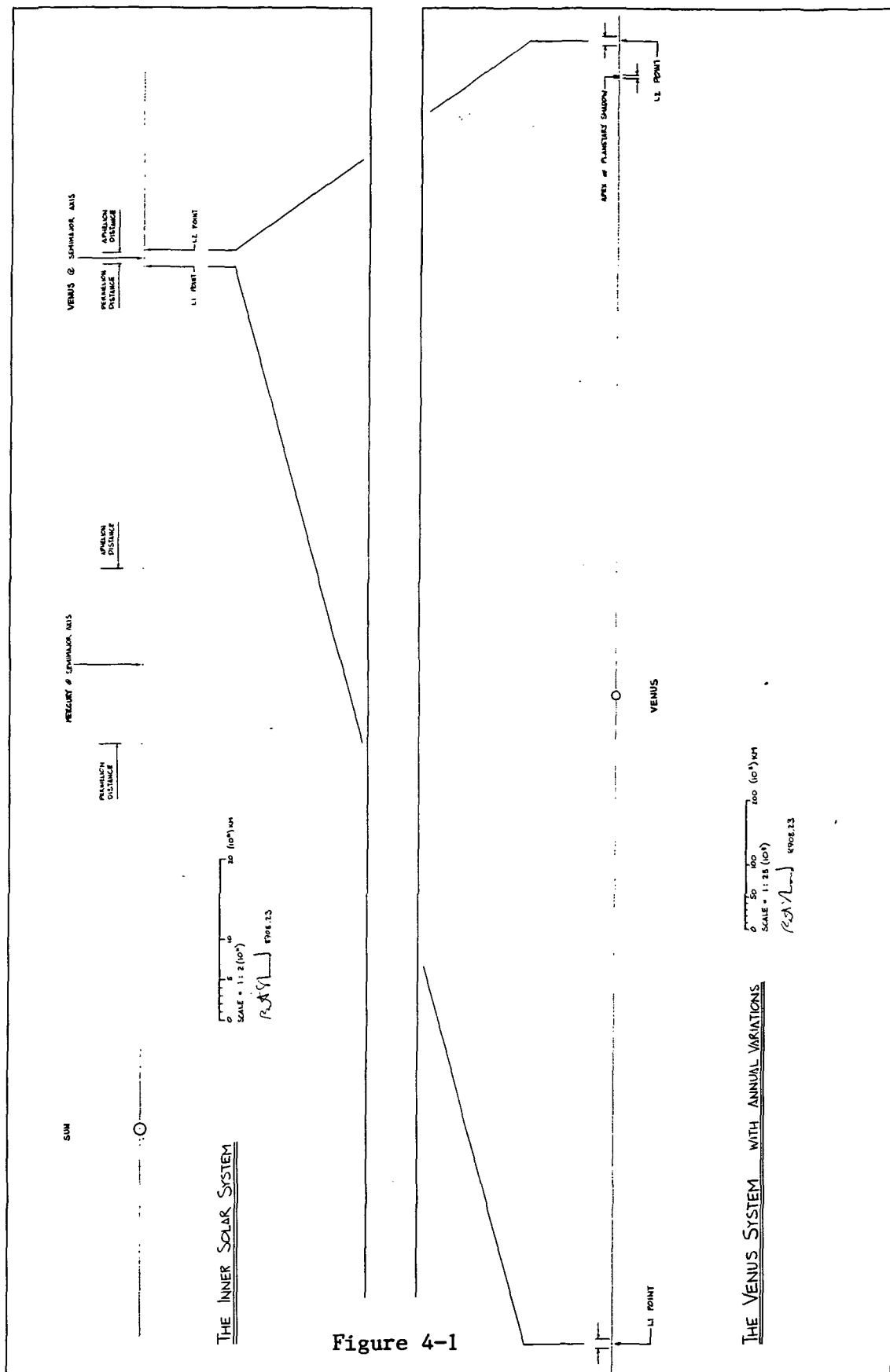


Figure 4-1

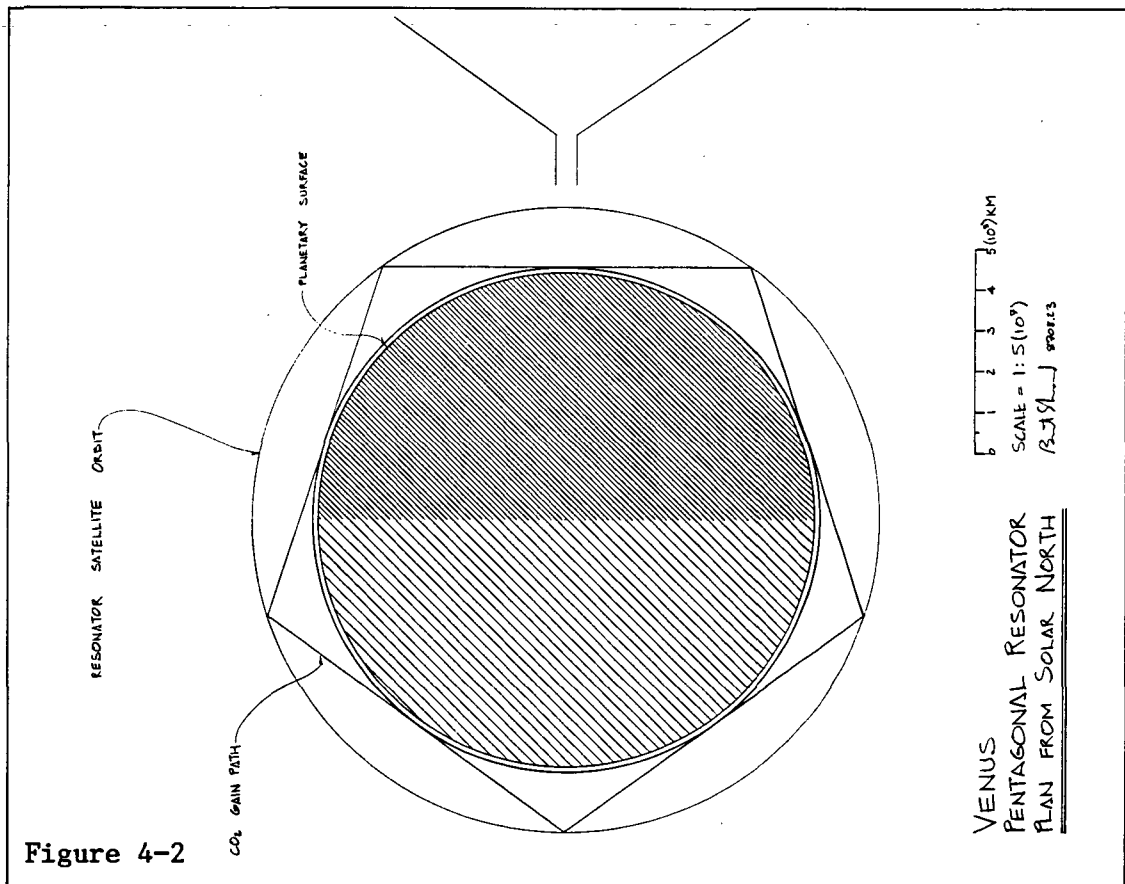
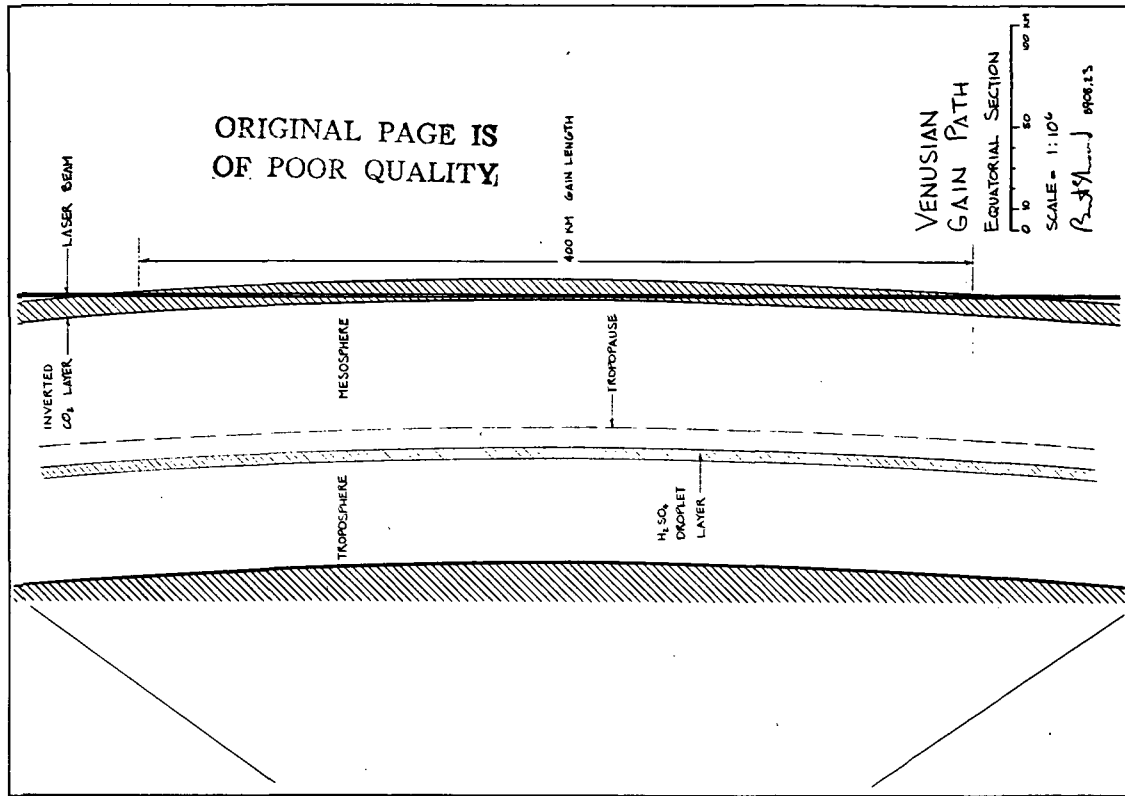
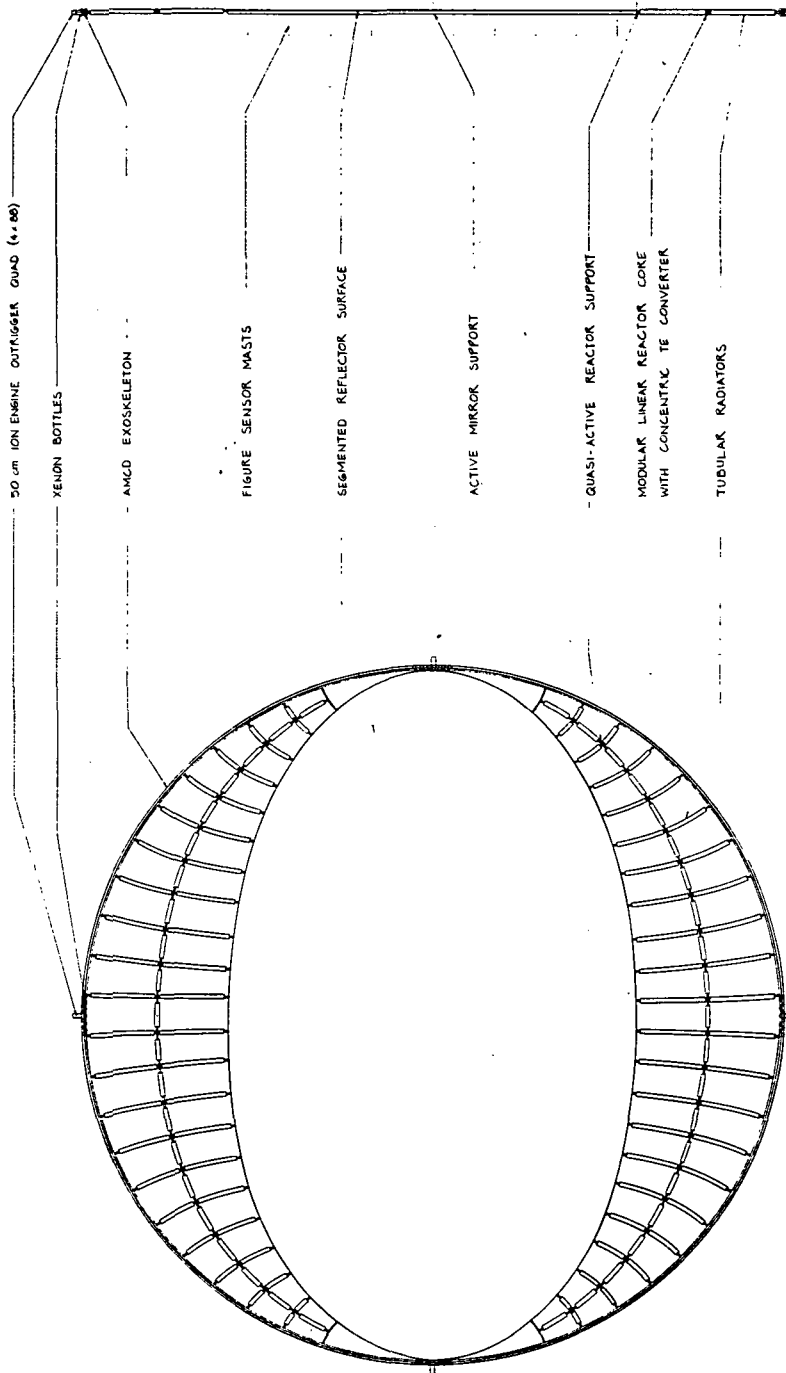


Figure 4-2

ORIGINAL PAGE IS
OF POOR QUALITY



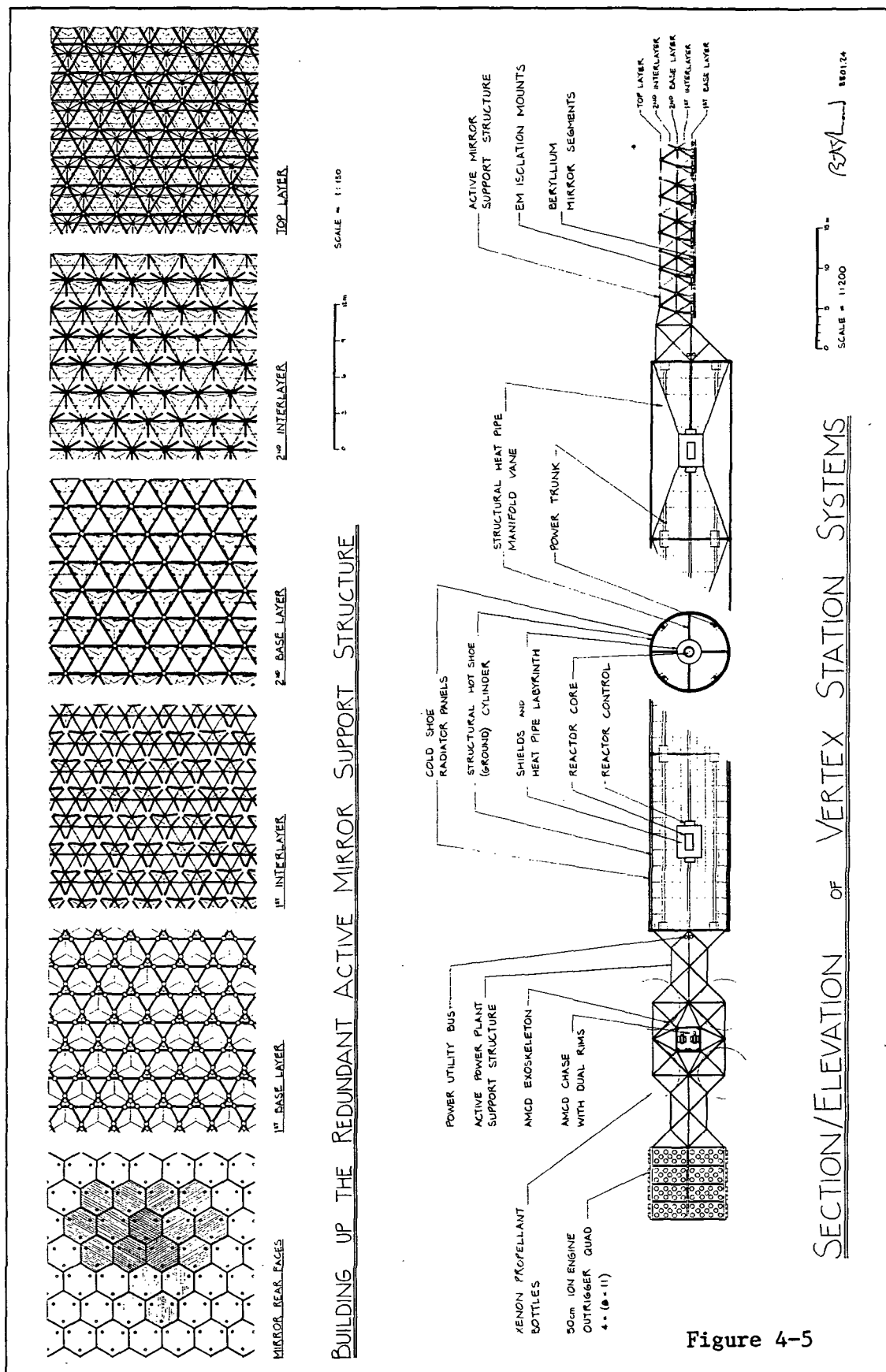
VENTRAL ELEVATION

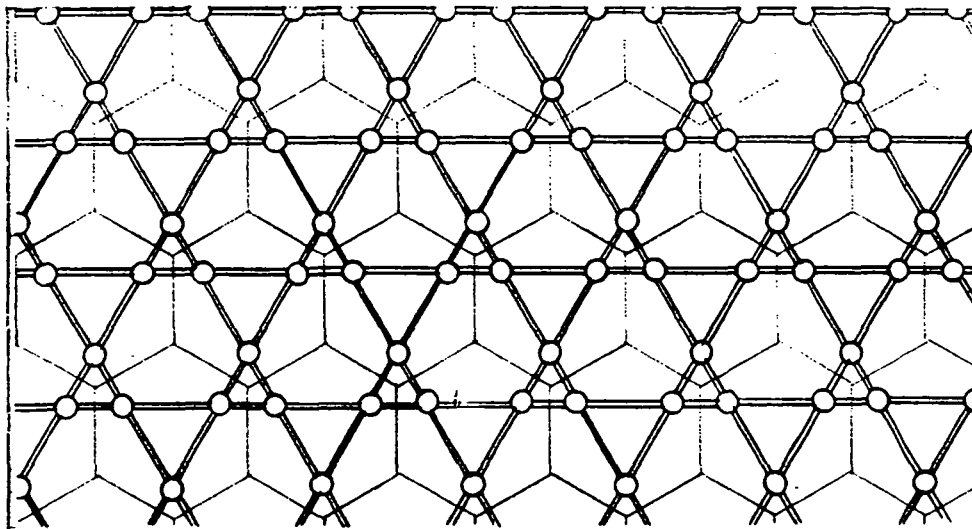
CROSS SECTION

0 50 100 200 m
SCALE = 1:5000
B. J. L. 10/1/20

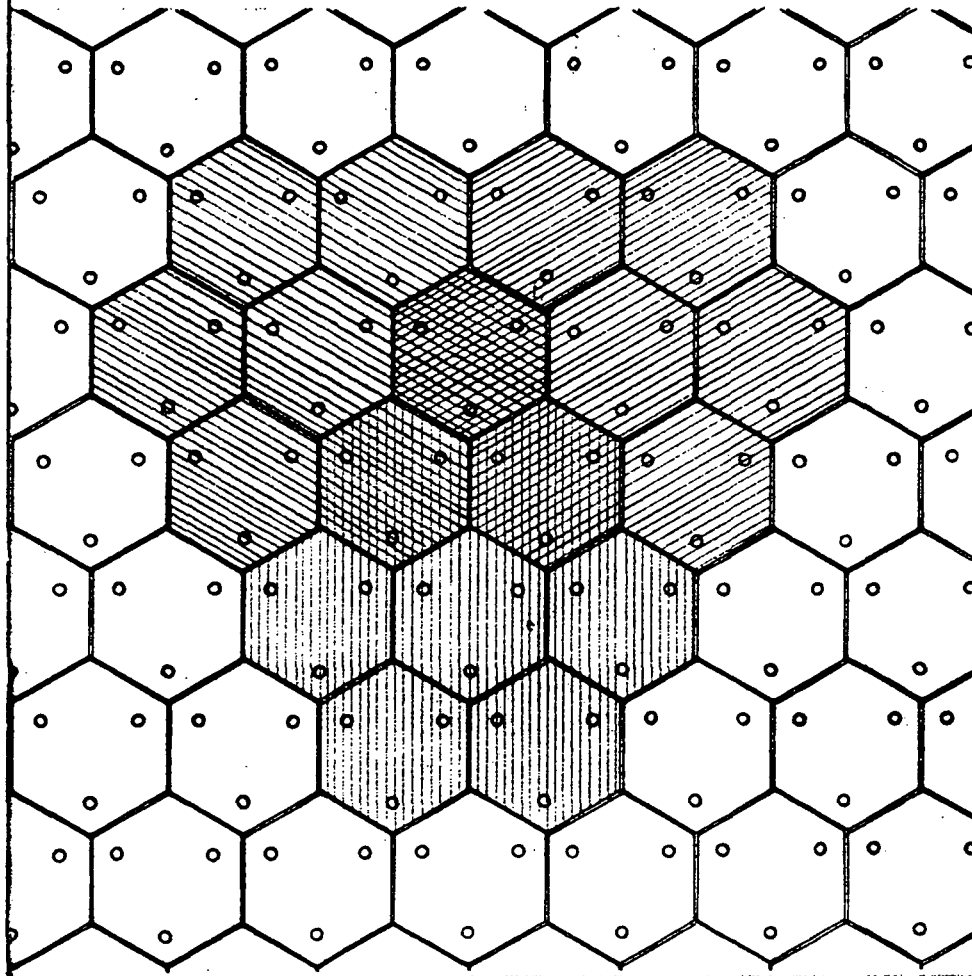
BASIC VERTEX STATIONS 2, 3, 4, 5

Figure 4-4





1ST BASE LAYER



MIRROR REAR FACES

Figure 4-6 Detail of Figure 4-5.

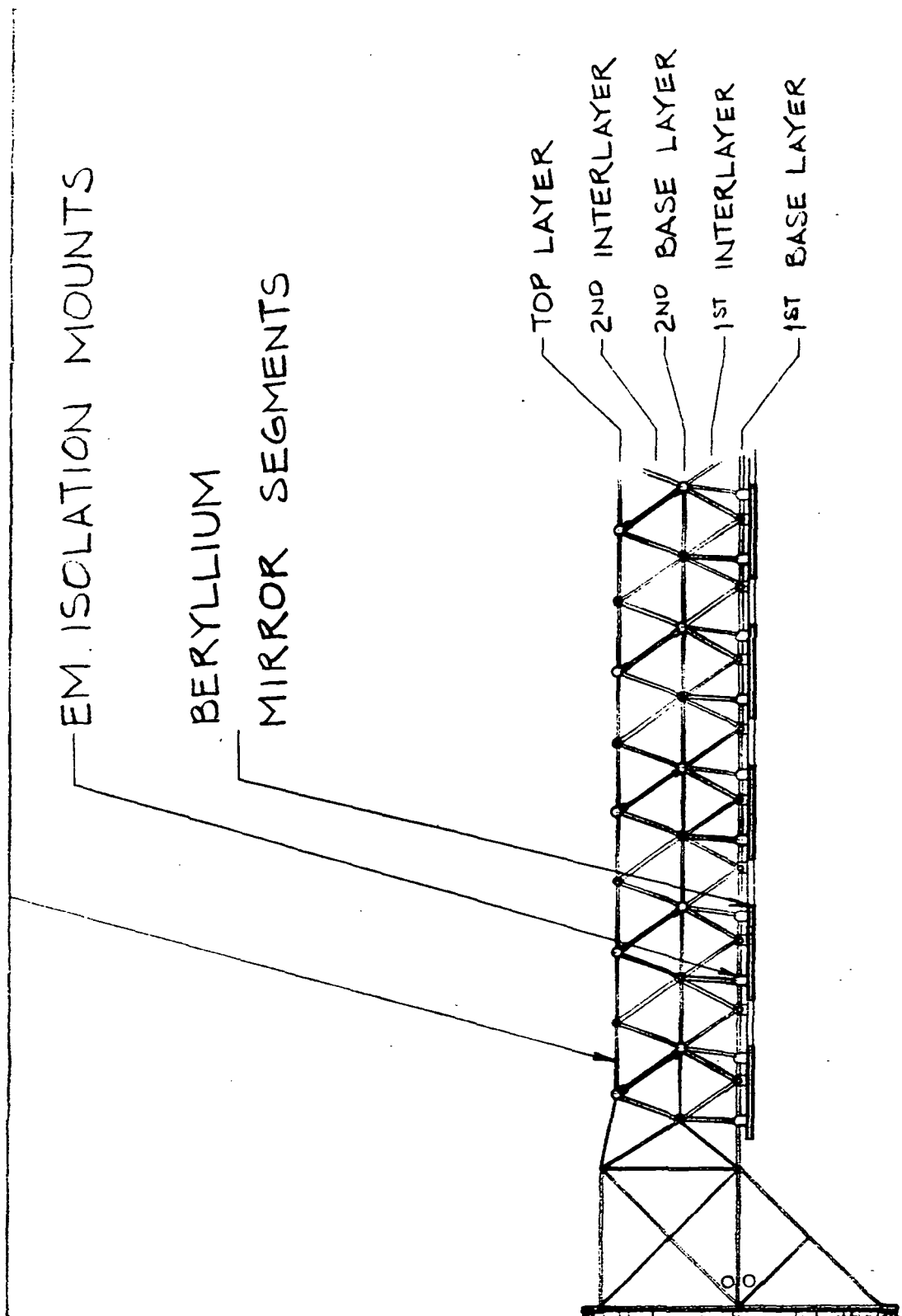
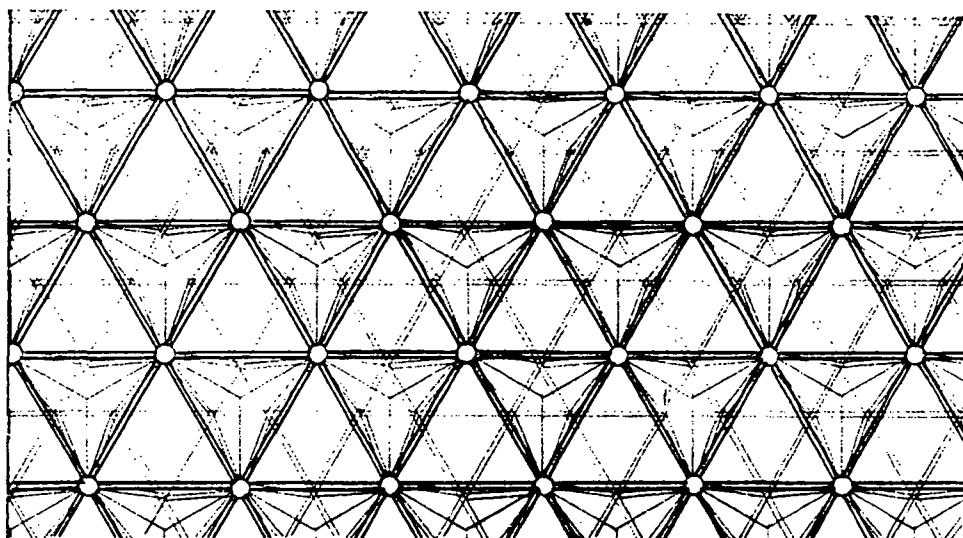
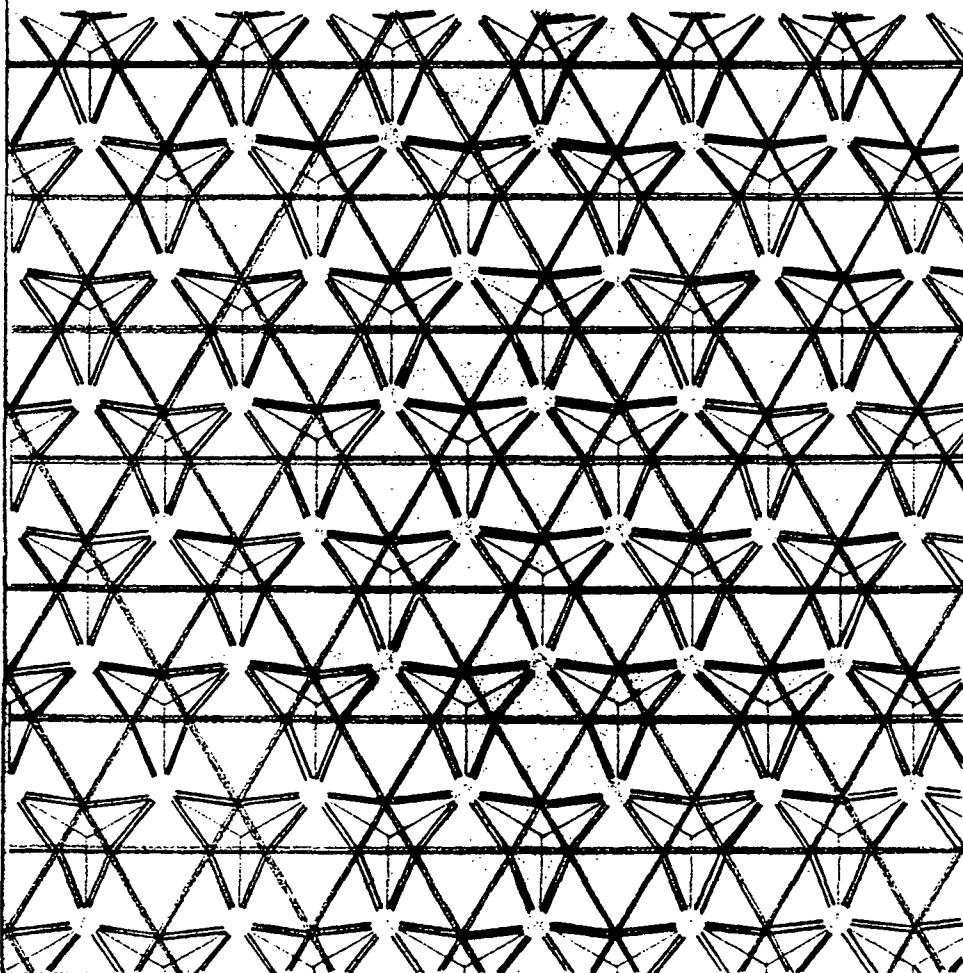


Figure 4-7 Detail of Figure 4-5.

ORIGINAL PAGE IS
OF POOR QUALITY

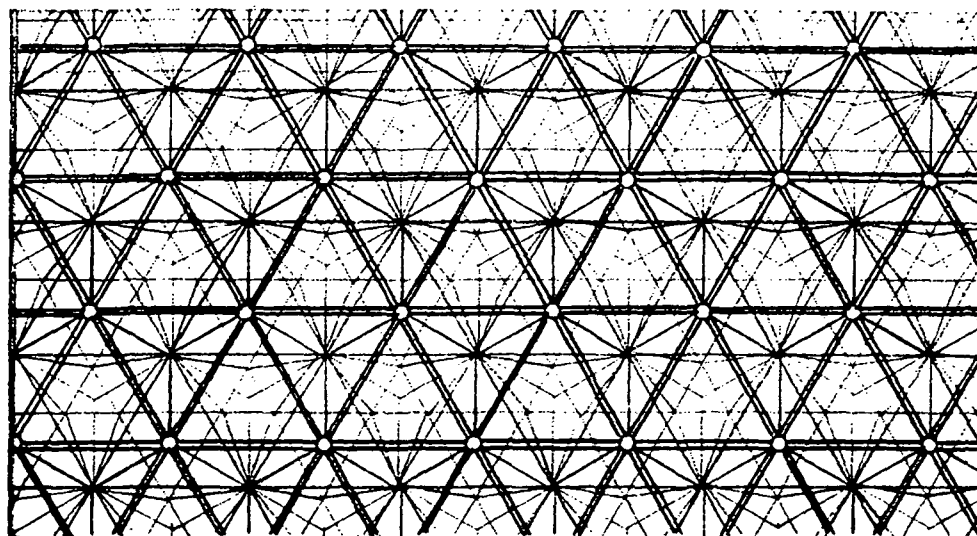


2ND BASE LAYER

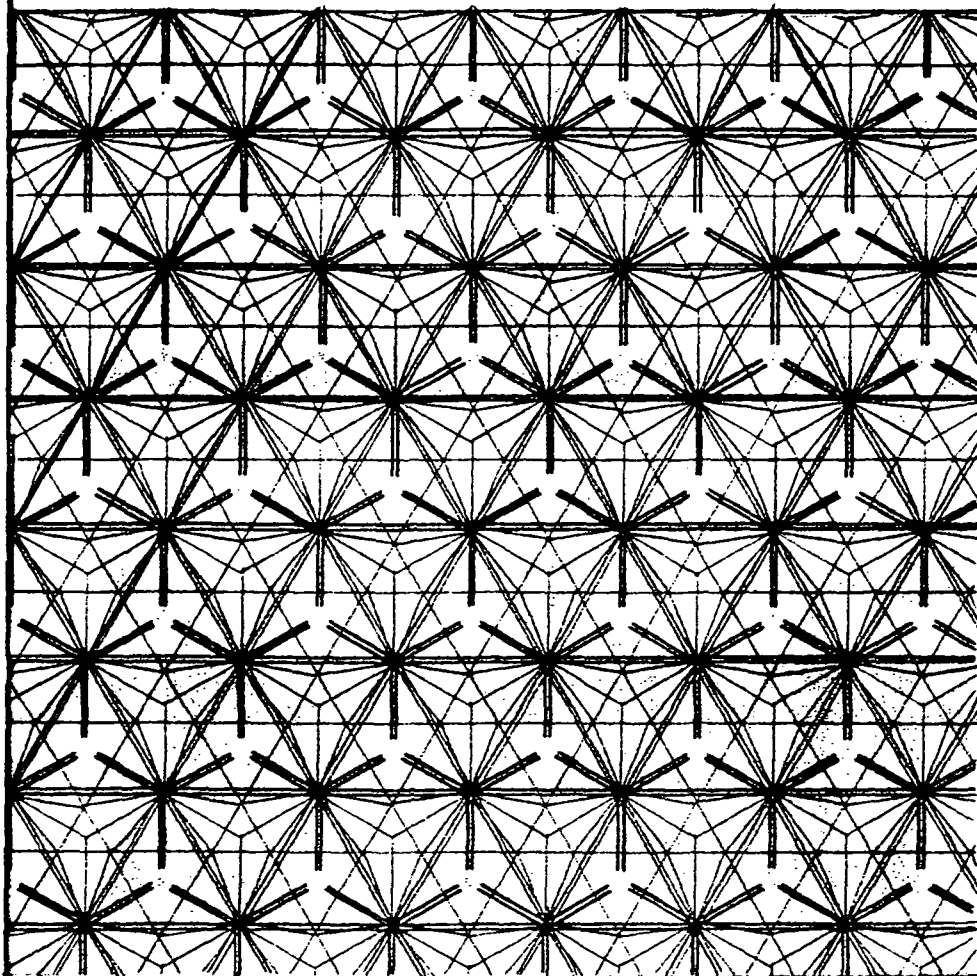


1ST INTERLAYER

Figure 4-8 Detail of Figure 4-5.



TOP LAYER



2ND INTERLAYER

Figure 4-9 Detail of Figure 4-5.

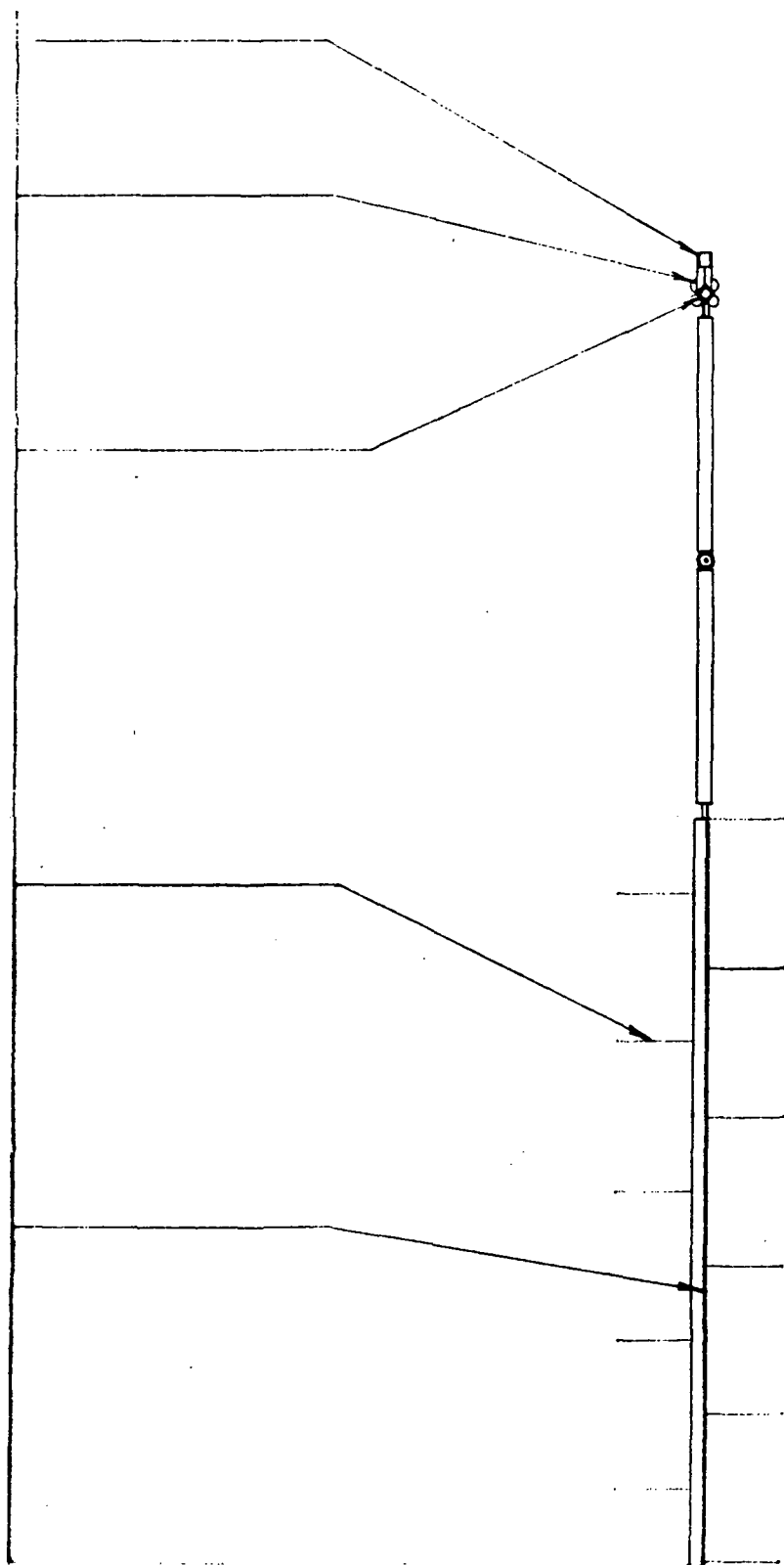


Figure 4-10 Detail of Figure 4-4.

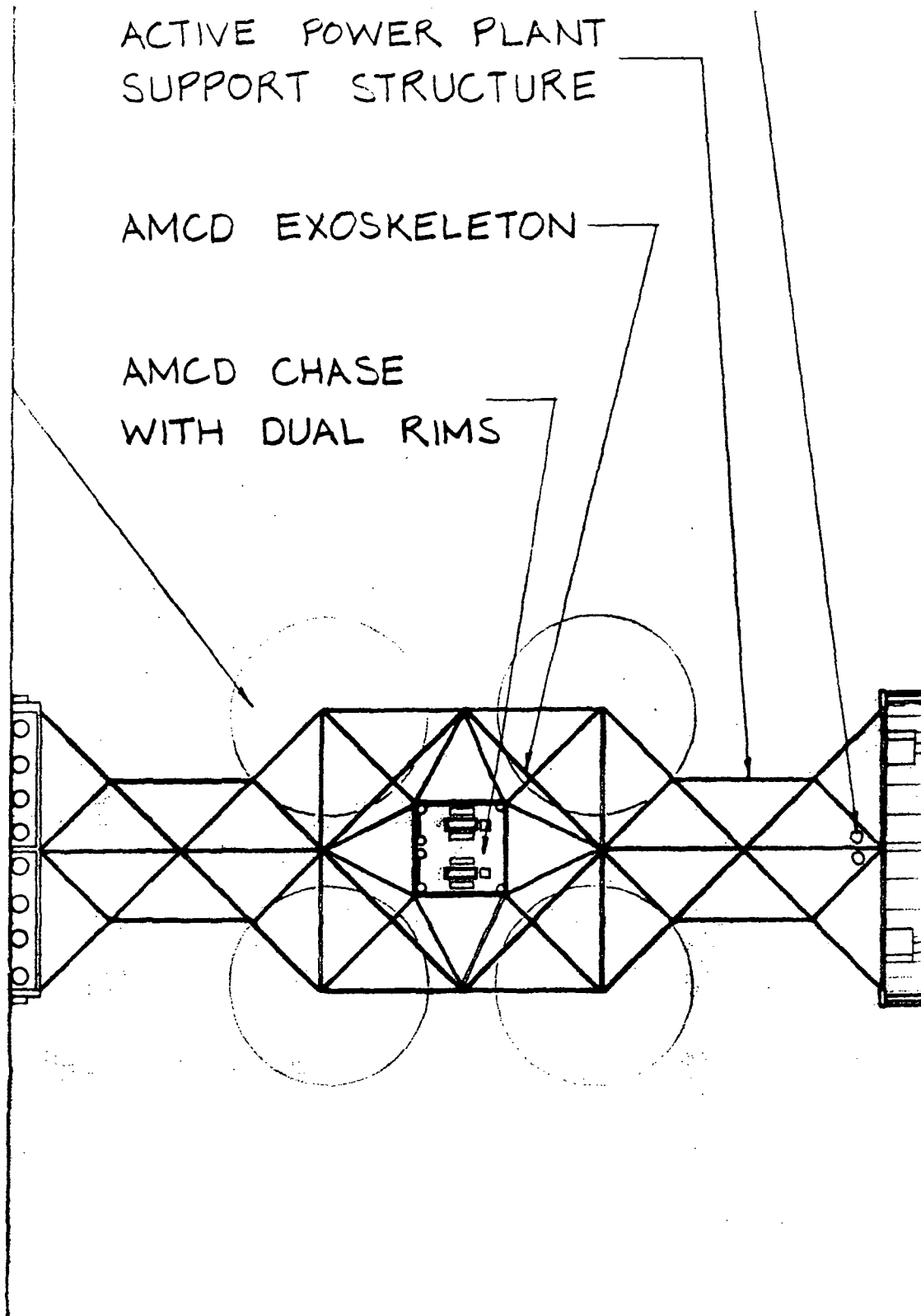


Figure 4-11 Detail of Figure 4-5.

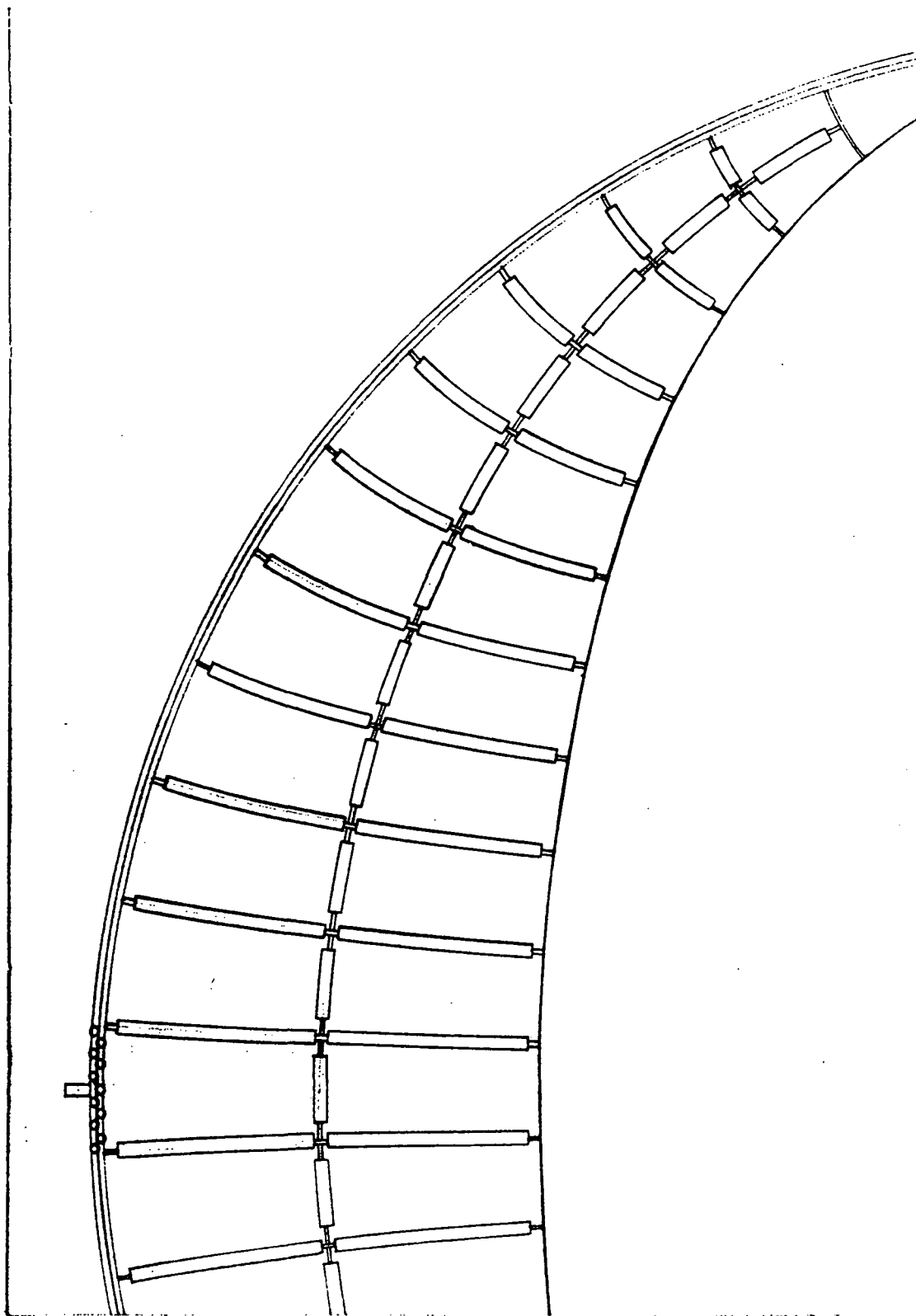


Figure 4-12 Detail of Figure 4-4.

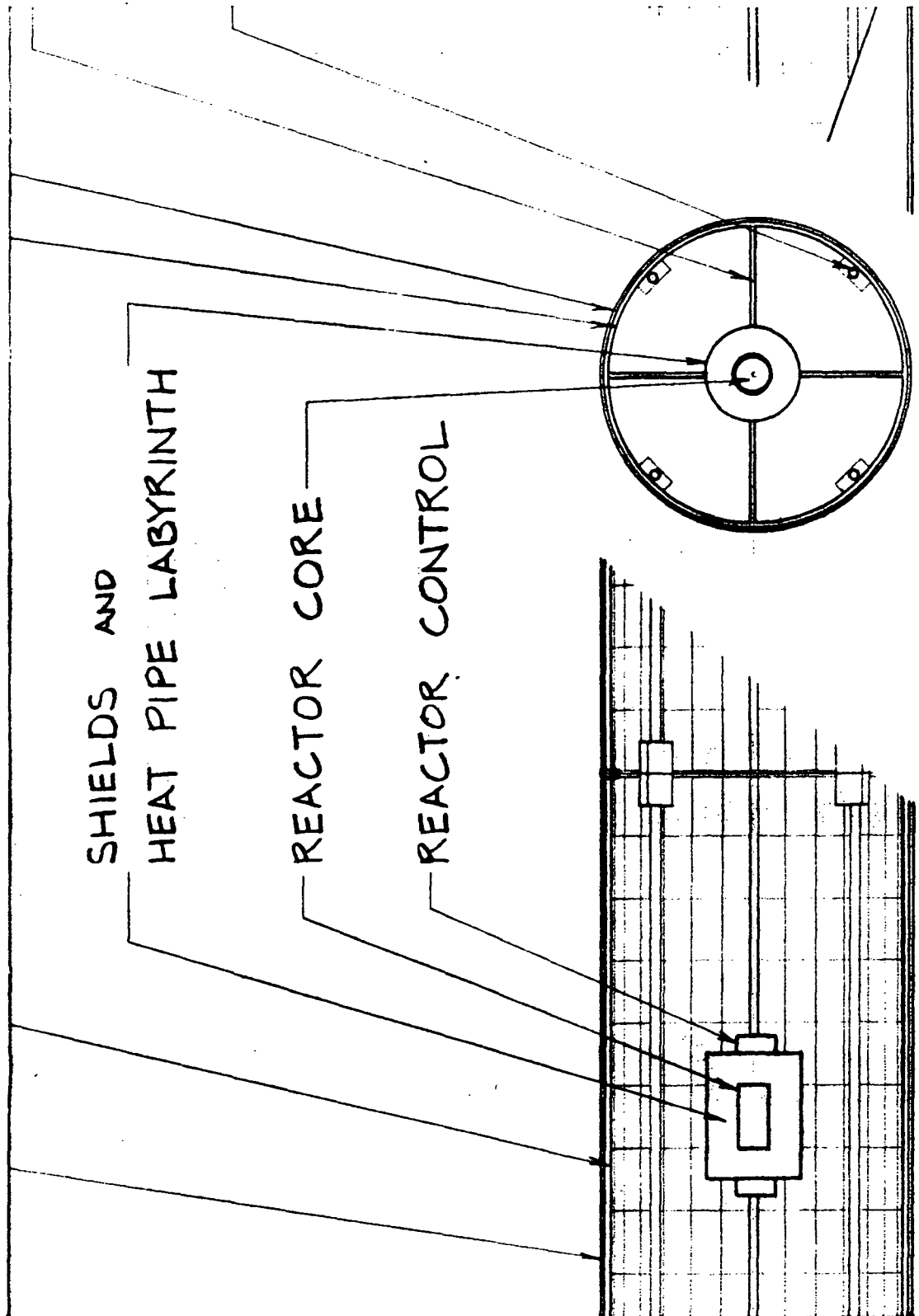


Figure 4-13 Detail of Figure 4-5.

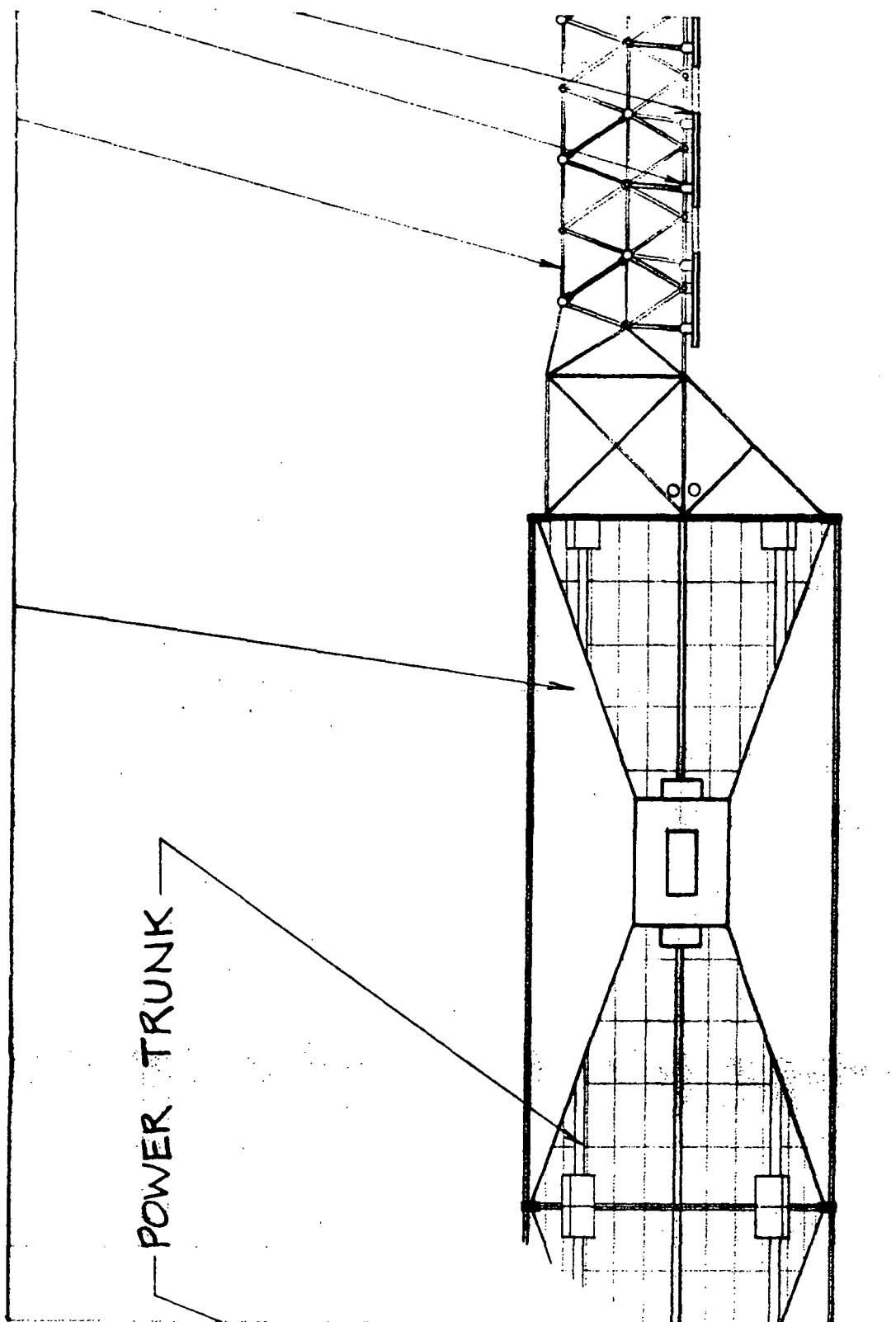


Figure 4-14 Detail of Figure 4-5.

XENON PROPELLANT
BOTTLES

50cm ION ENGINE
OUTRIGGER QUAD
 $4 \times (8 \times 11)$

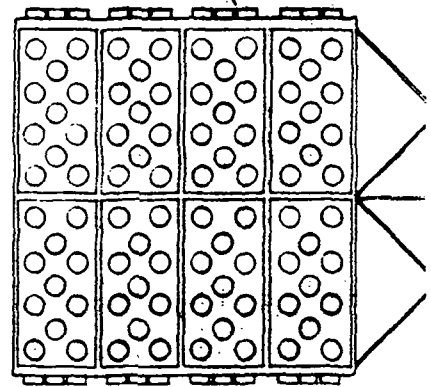


Figure 4-15 Detail of Figure 4-5.

THE $(\beta-\gamma-\delta-\epsilon)$ MIRROR CONSTELLATION IS
 ROTATED 90° ABOUT THE 1-2 BEAM AXIS
 TO SHOW THE NOMINAL γ POSITION 1KM
 BELOW THE ORBIT PLANE, REQUIRED BY
 CONSISTENT POLARIZATION ORIENTATION
 OF THE β DIFFRACTION GRATING.
 THE δ MIRROR IS BEHIND β IN
 THIS VIEW.

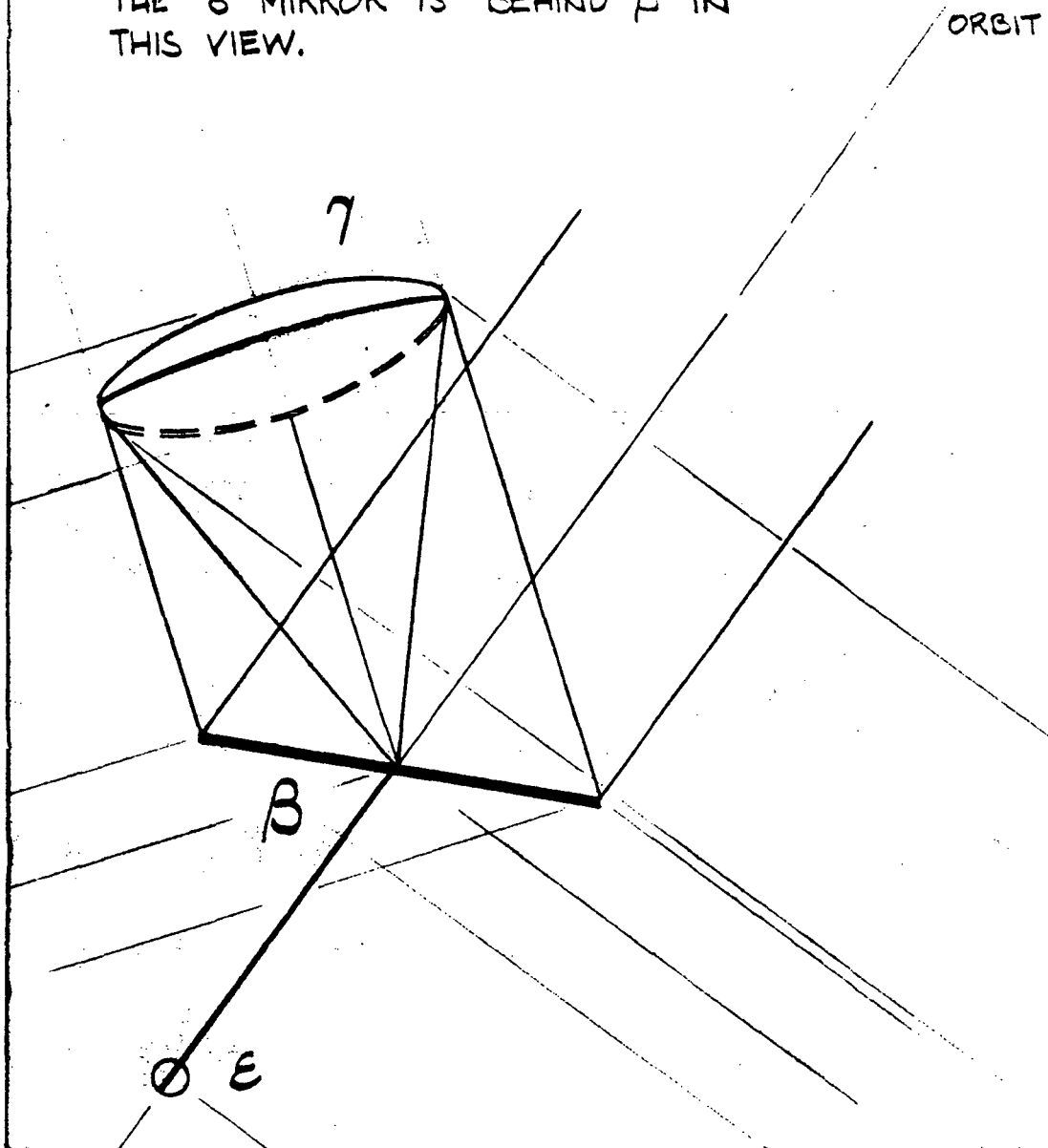
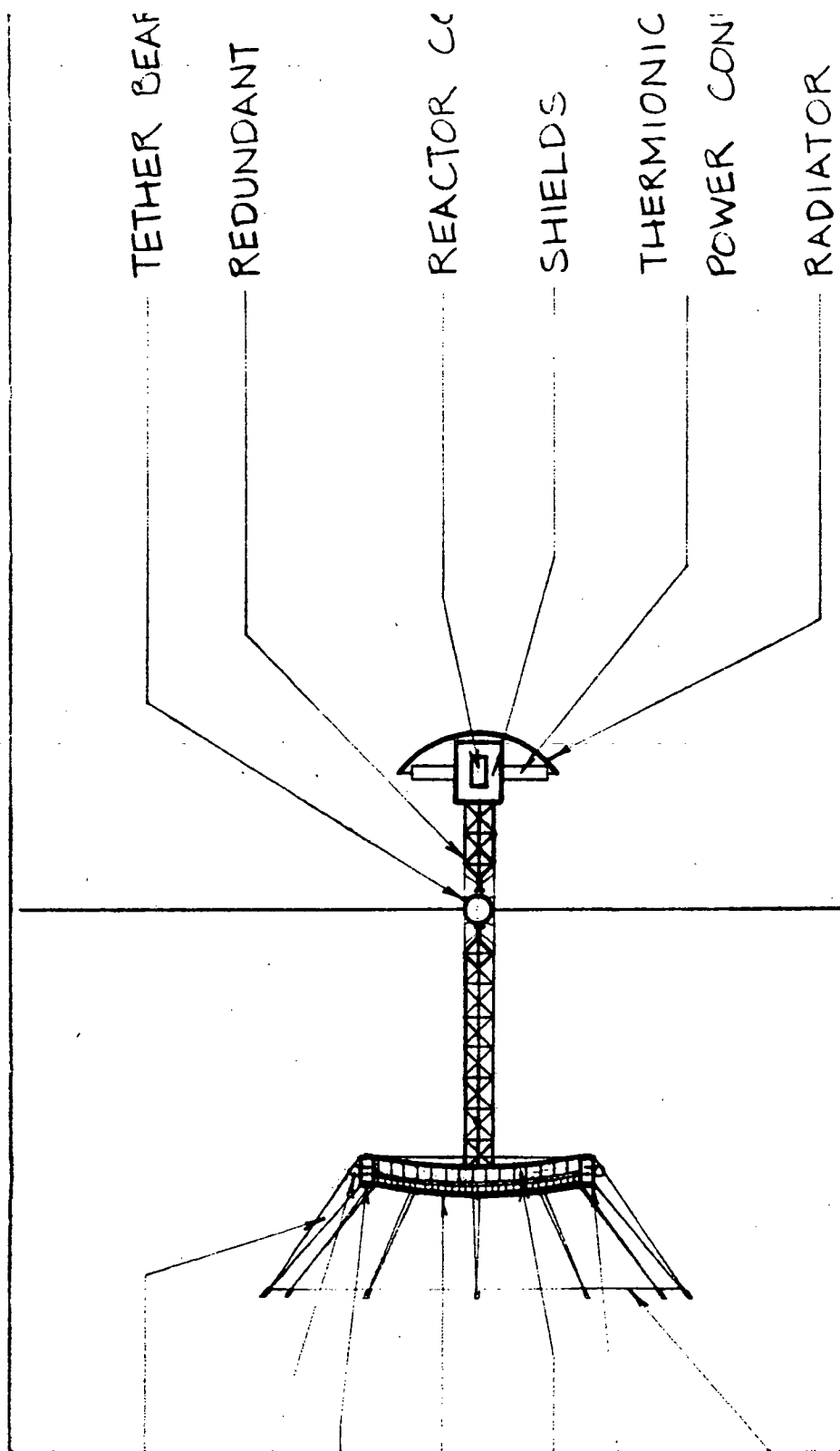


Figure 4-16 Detail of Figure 4-3.



18 LONG SECTION

Figure 4-18 Detail of Figure 4-17.

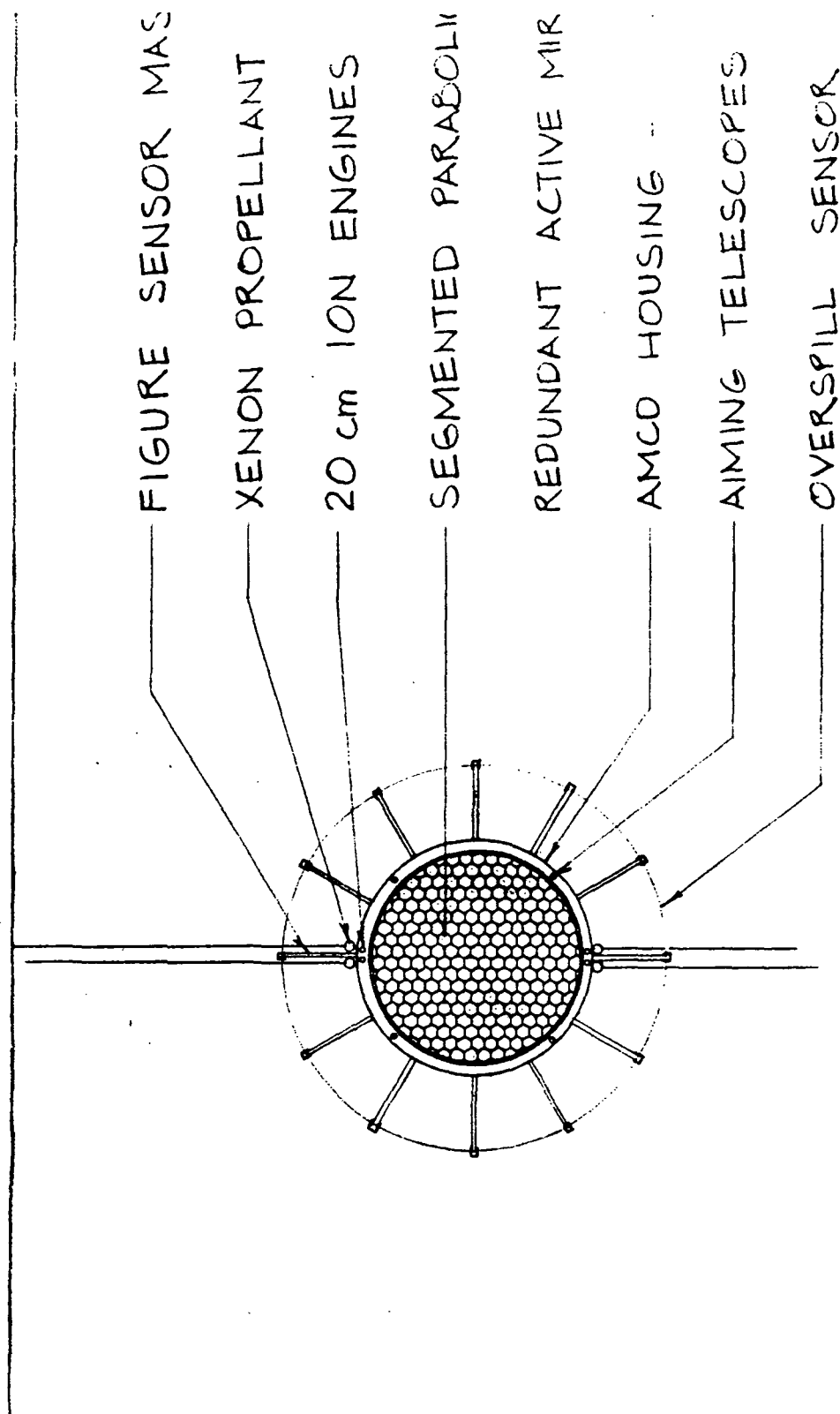


Figure 4-19 Detail of Figure 4-17.

18 FRONT ELEVATION

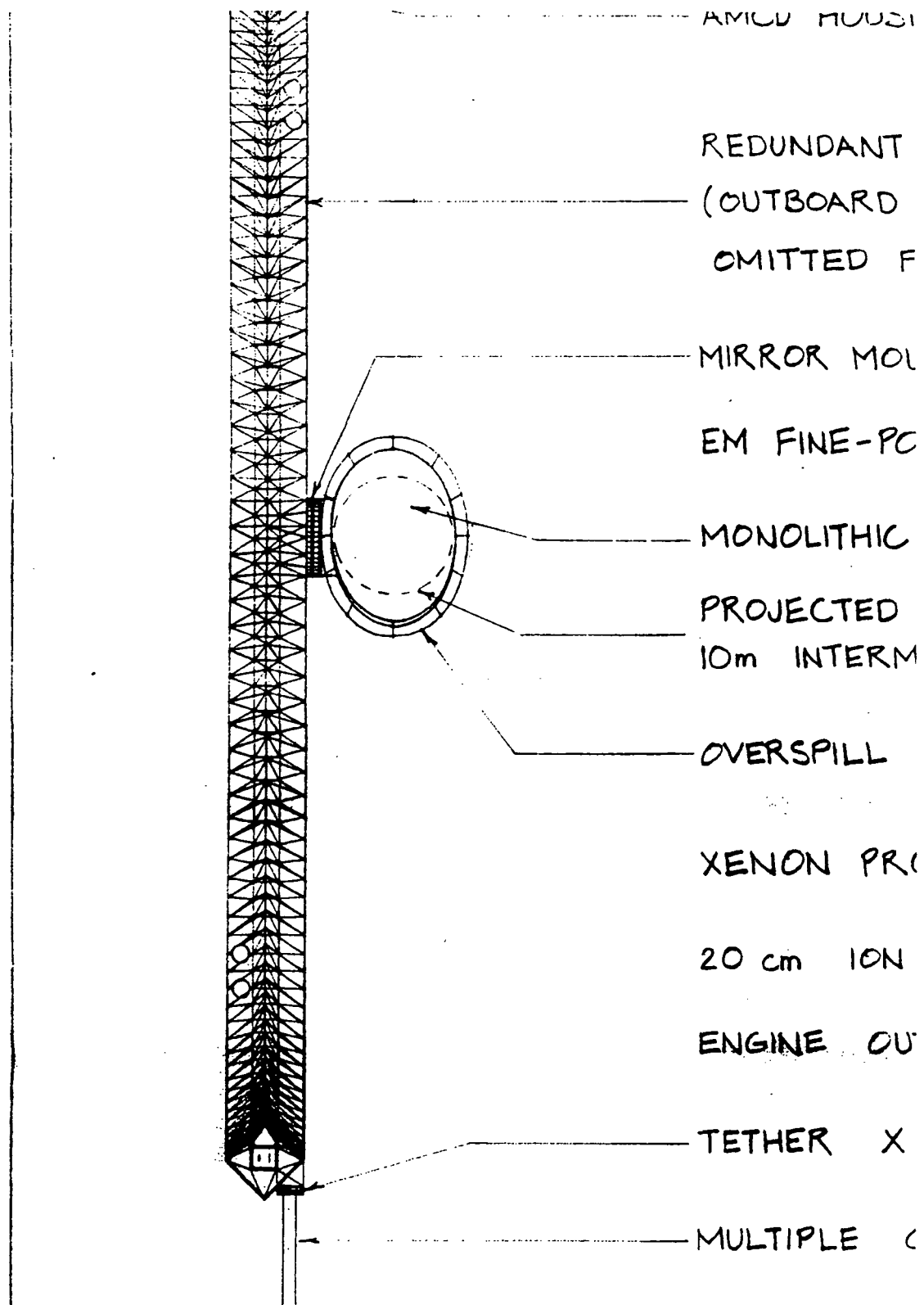


Figure 4-20 Detail of Figure 4-17.

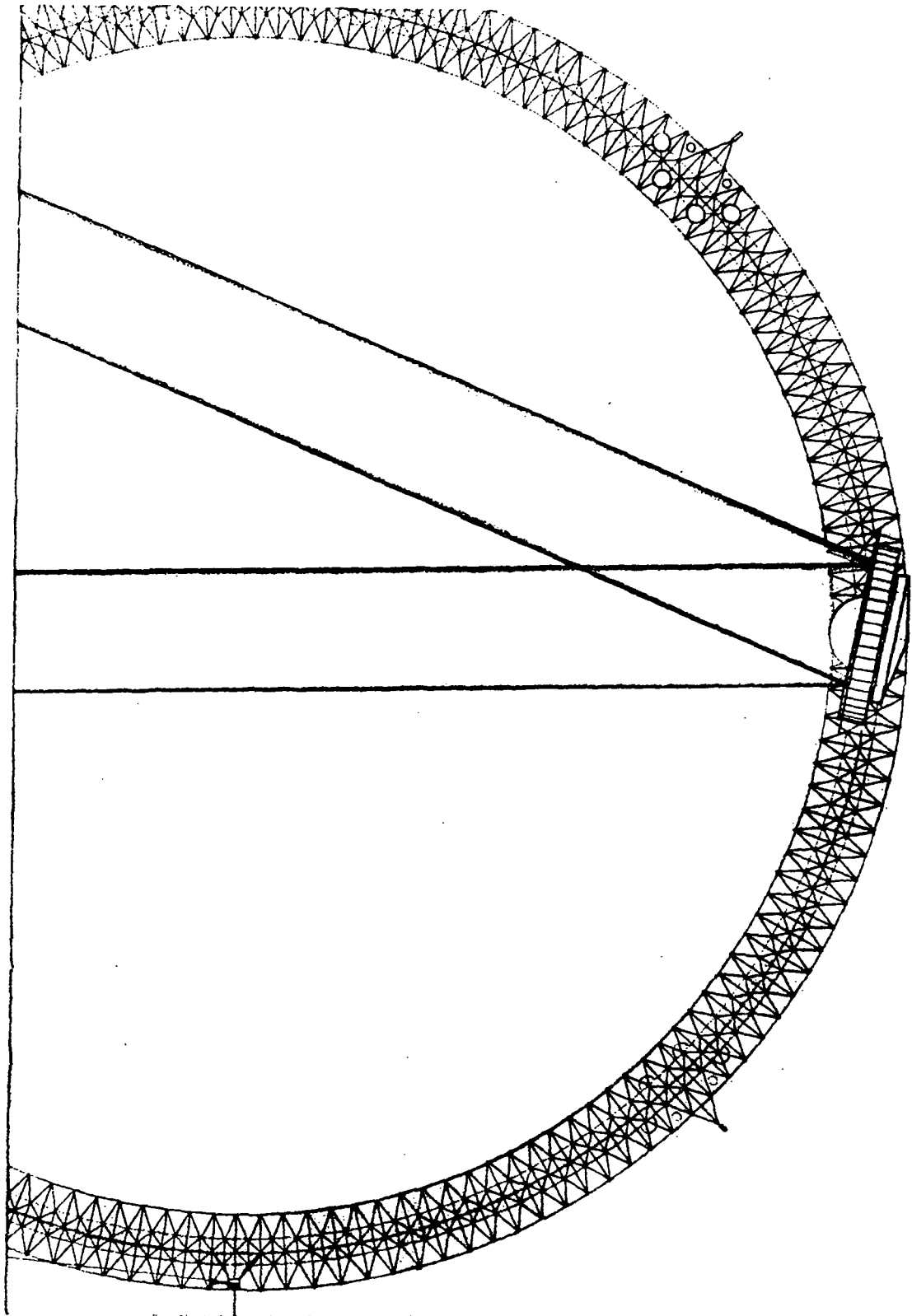


Figure 4-21 Detail of Figure 4-17.

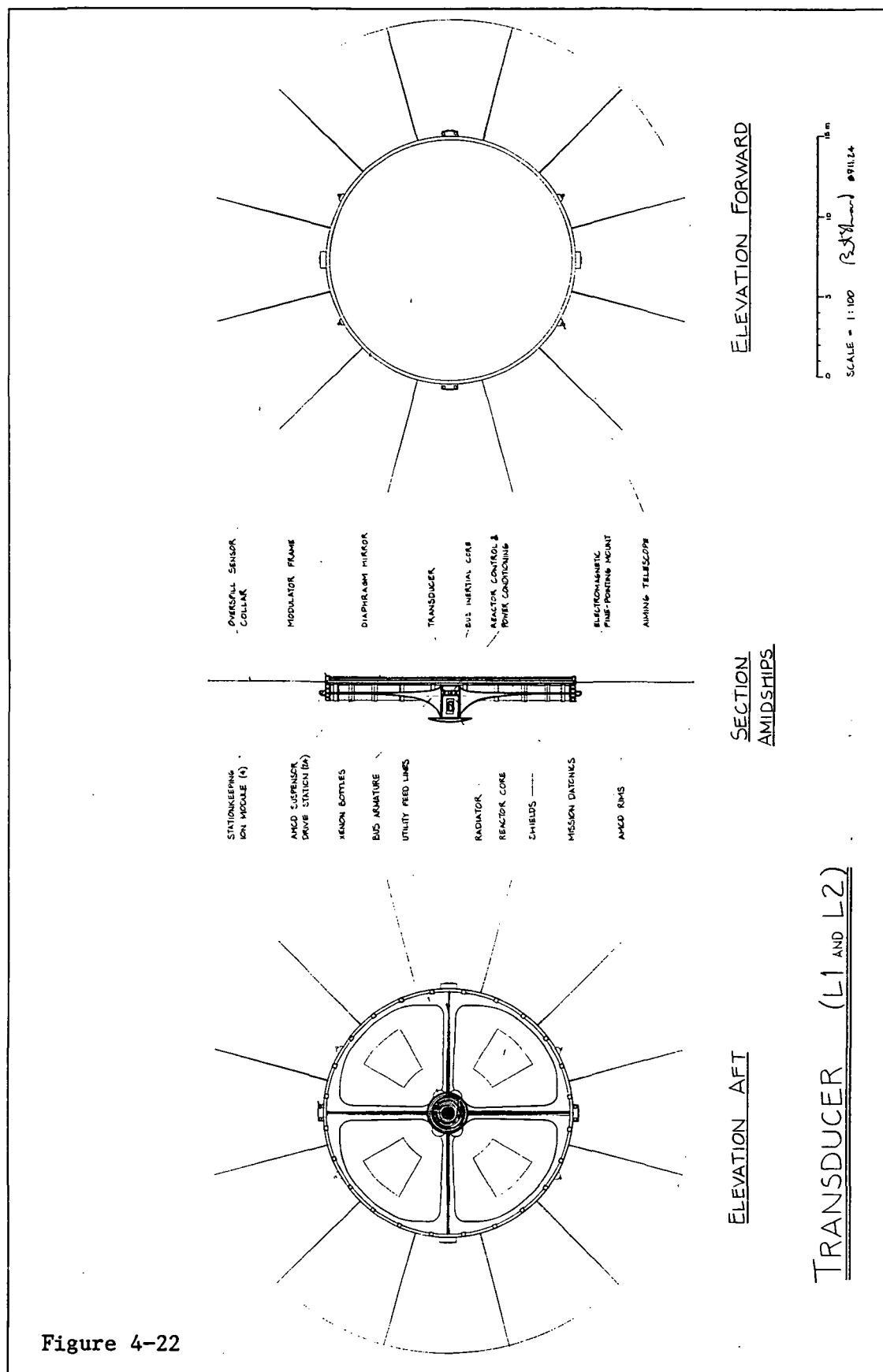


Figure 4-22

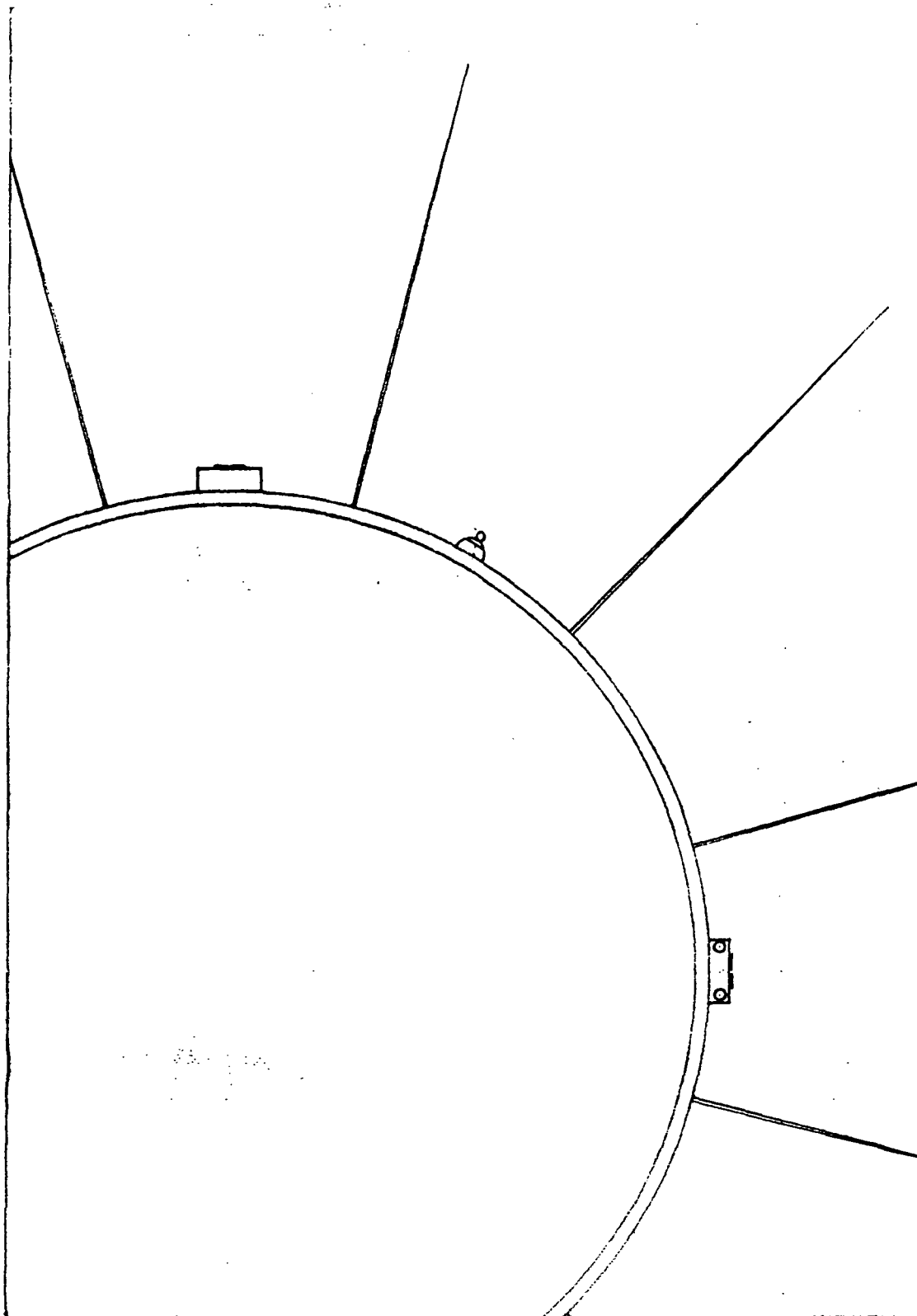


Figure 4-23 Detail of Figure 4-22.

ORIGINAL PAGE IS
OF POOR QUALITY

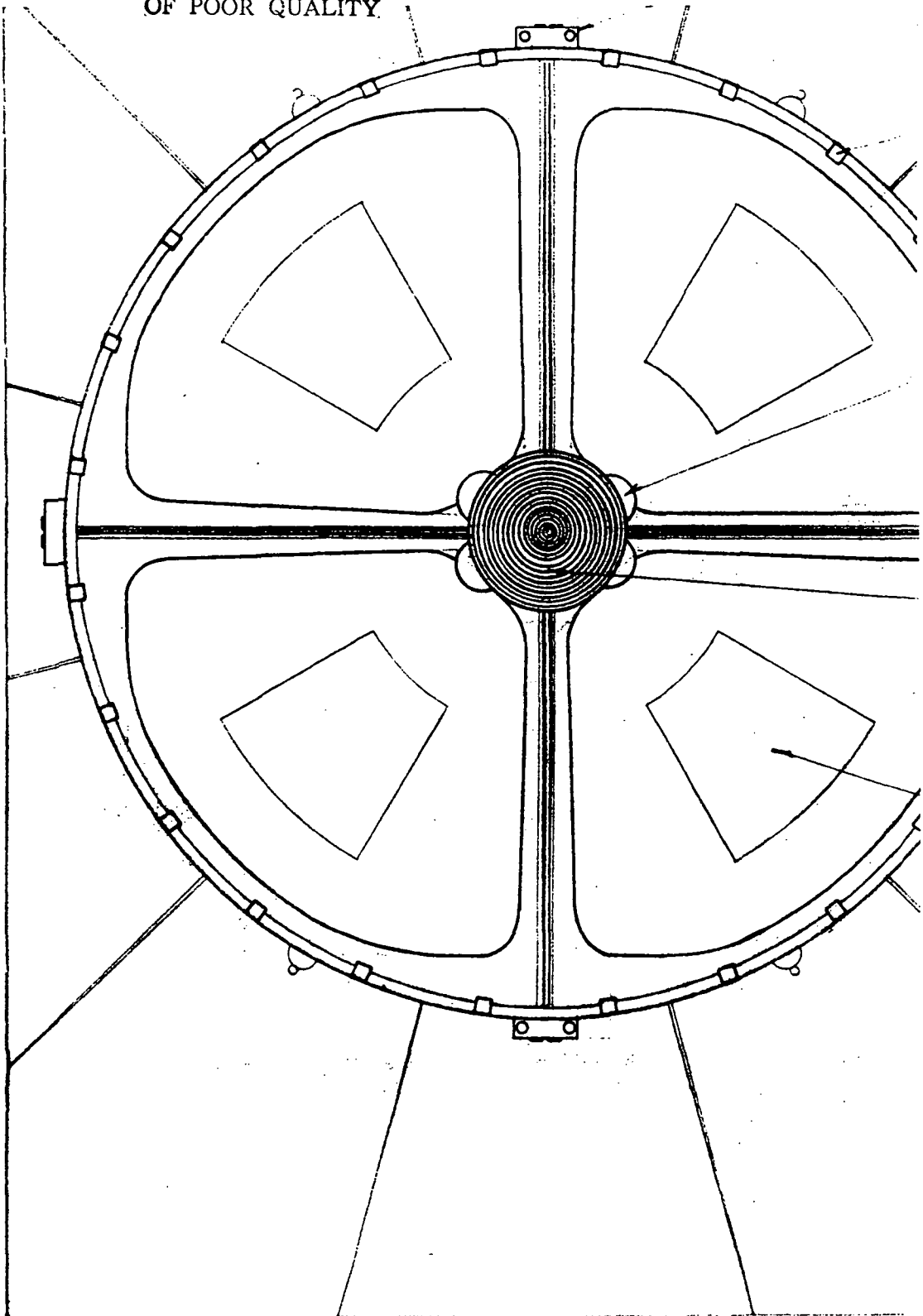


Figure 4-24 Detail of Figure 4-22.

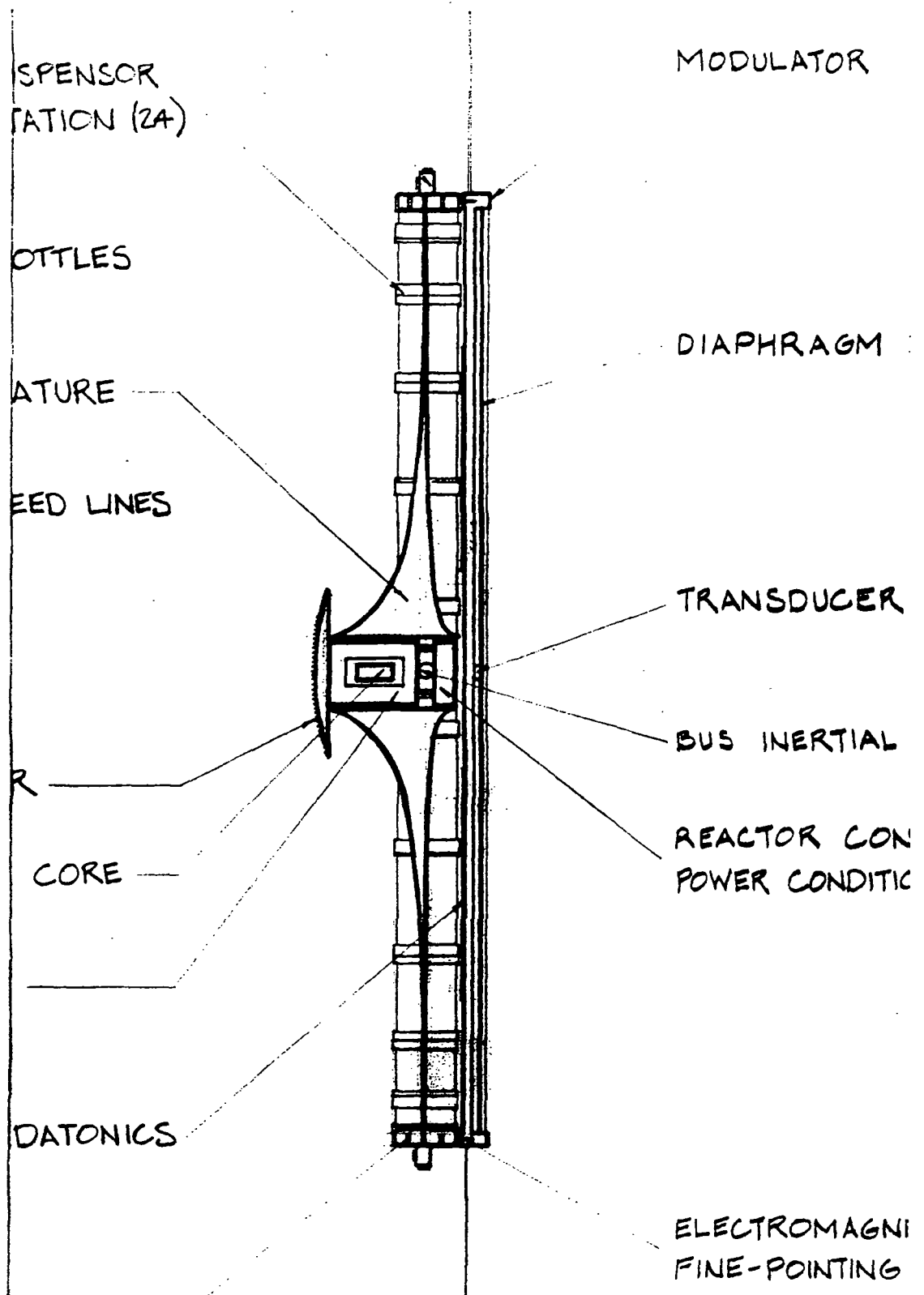


Figure 4-25 Detail of Figure 4-22.

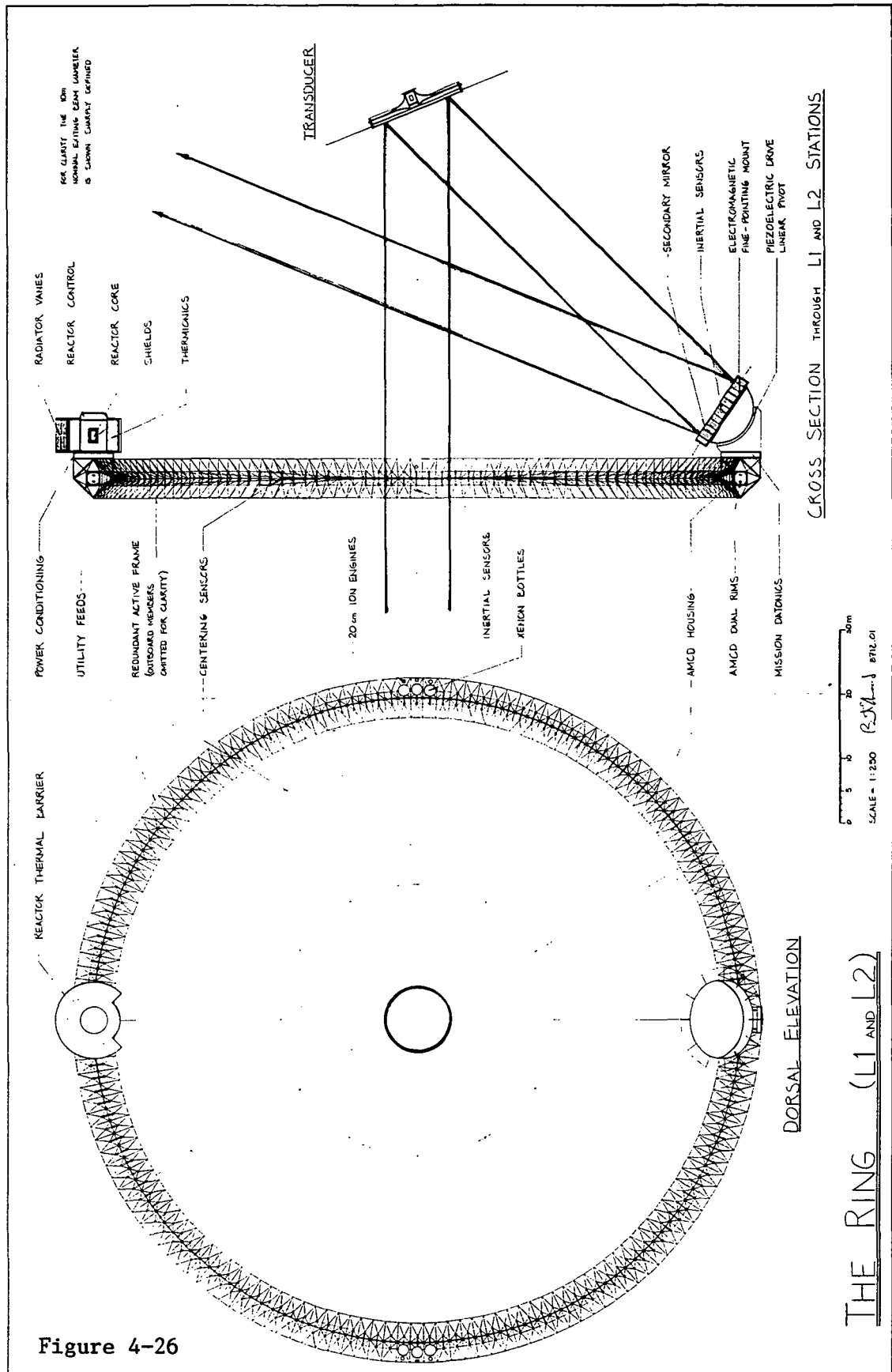


Figure 4-26

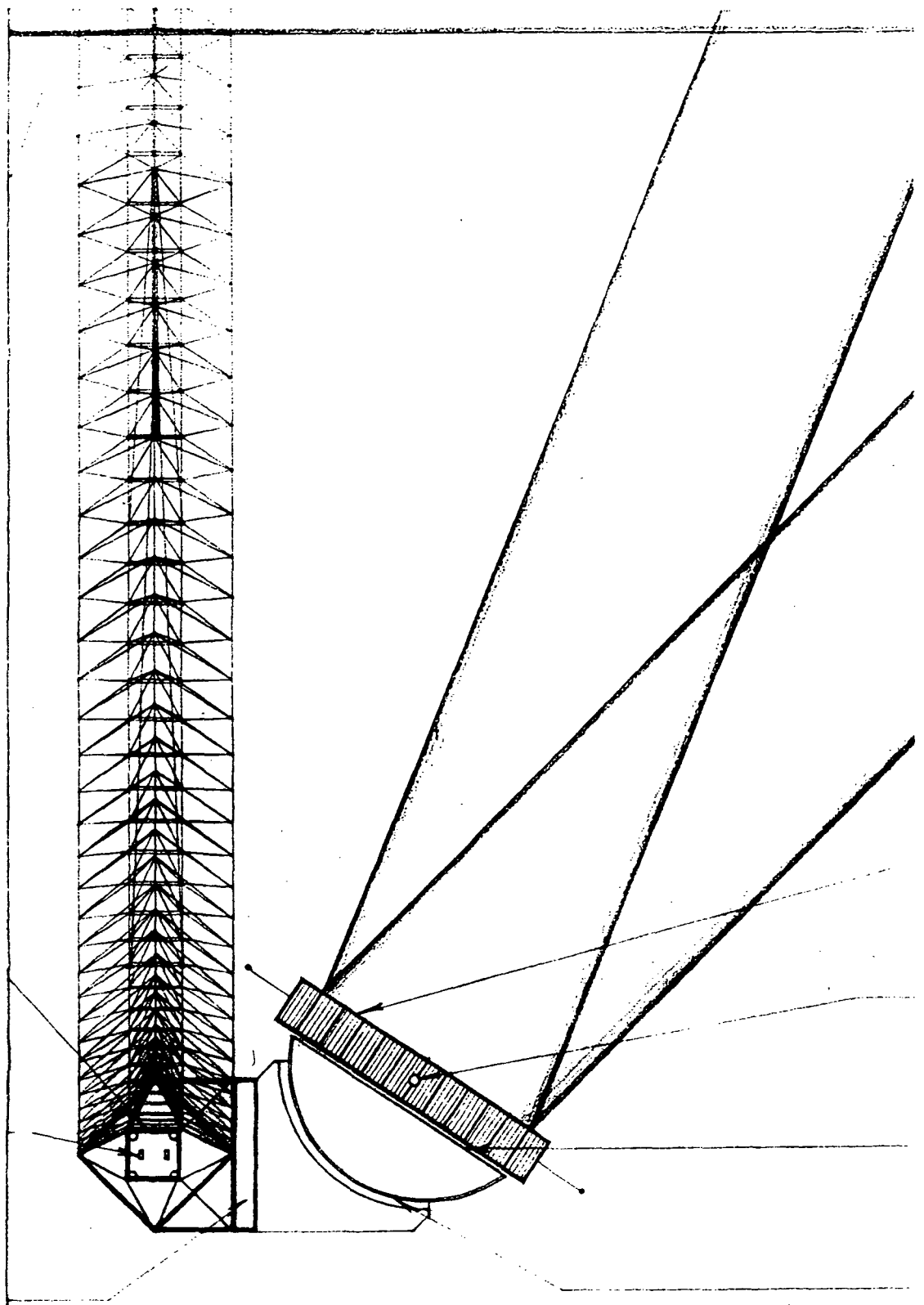


Figure 4-27 Detail of Figure 4-26.

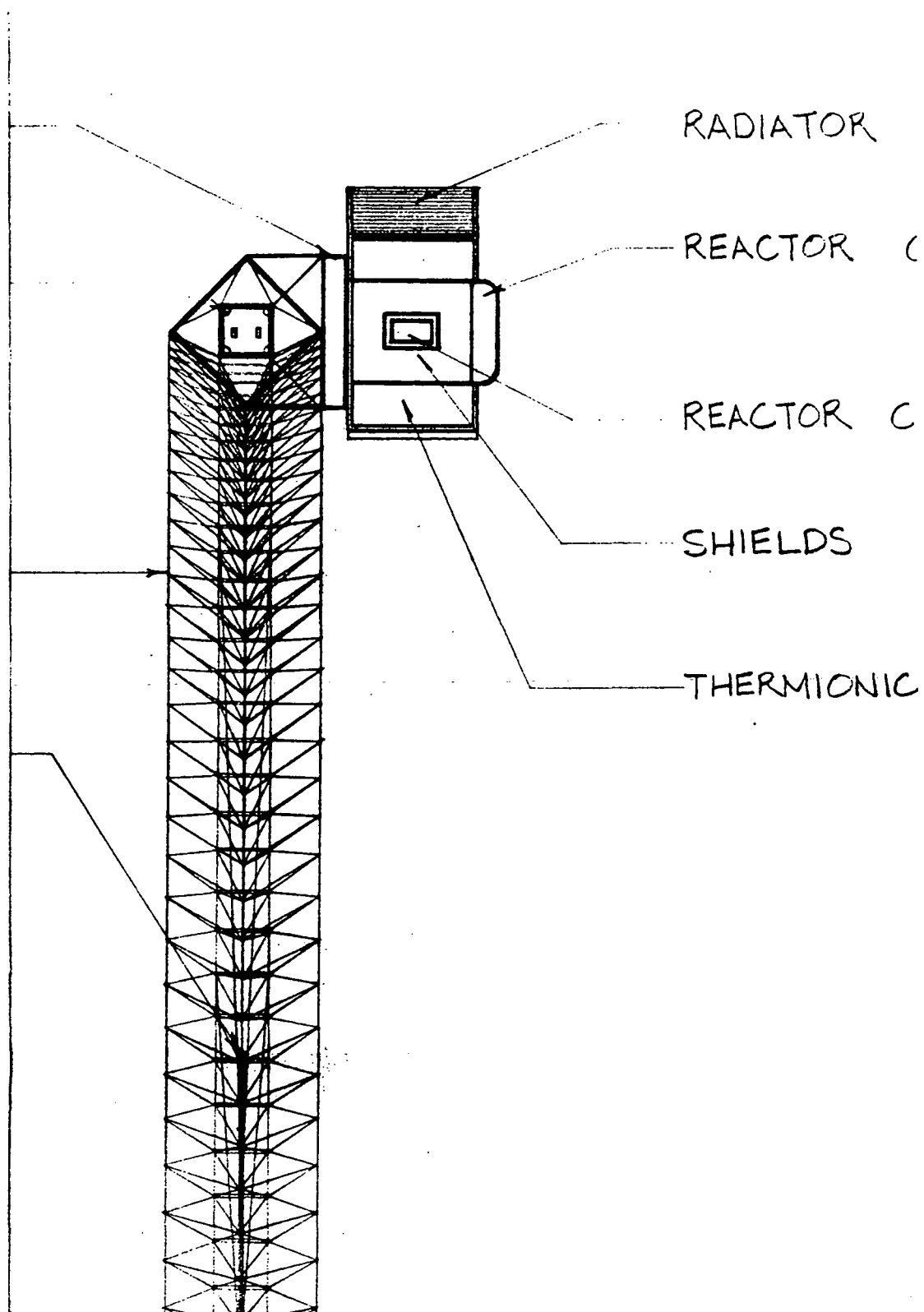


Figure 4-28 Detail of Figure 4-26.

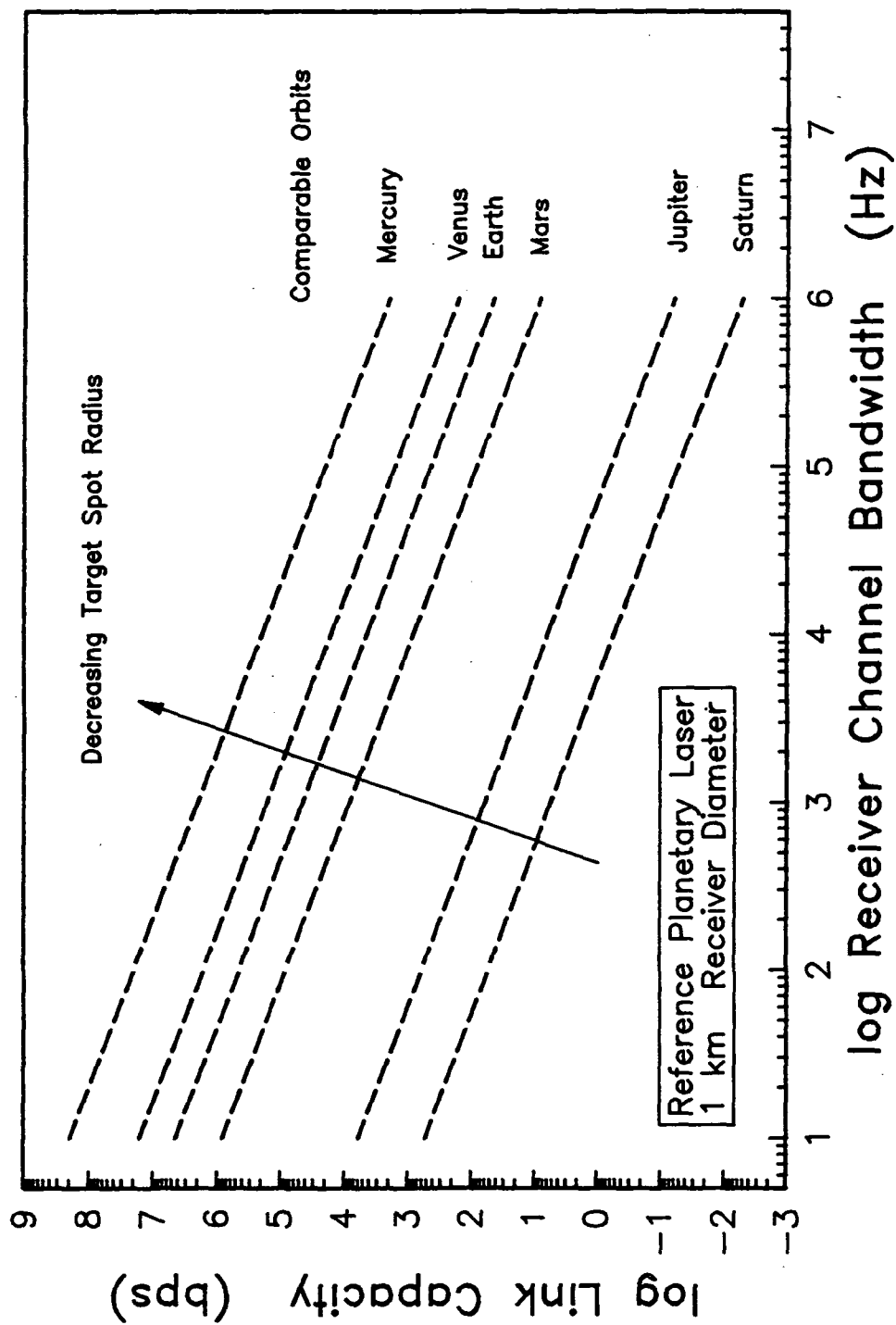


Figure 4-29 Link performance for the reference CETI mission; interstellar data transfer rate out to 82 ly.

Appendix A4-1 Basic Vertex Station (2,3,4,5) systems sketch.

Power Requirement

Mirror Control - Since the 180 kg mirror segments need to be actuated quickly, a power price must be paid. Based on the EM actuators referenced in Chapter 8, we choose a not-too-conservative 50 W per mount. Assuming this is an average value for their cycling operation between positioning and holding, we multiply it by the number of actuators per mirror segment (3) and then by the number of mirror segments surfacing the 1 km x 1.7 km elliptical reflector (230,000).

35 MW

Mirror Support Structure - The backup structure is a two-layer redundant truss built up of active members and nodes. Appendix A4-7 specifies 20 W average power to run such members, and Appendix A4-3 calls for 10 W per member for short members. We can allocate 21 members through the depth of the mirror support truss to each segment, 18 of which are at least 2 m long, and 3 of which are about 1 m long. Thus we budget $(18(20) + 3(10))230,000 = 90$ MW. Because of the large number of members meeting at each active node, we increase this total by 50 % to account for them.

135 MW

AMCD Exoskeleton - This active structure is completely analogous to those developed for 1c and the Ring (Appendices A4-3 and A4-7), but larger in two ways. Its circumference and its enclosed chase are both larger. Because of the linear nature of both the circle and the truss elements, both increases can be used to scale up the power linearly from the value used in those accountings: $102(850/50)(3/2)$

3 MW

Reactor Support Structure - We refer to this structure system as quasi-active, since only a tiny fraction of it is controlled actively. Although many fewer members are involved than for the AMCD exoskeleton, we expect each member to use more power because of the relatively larger job it must do.

2 MW

AMCD Control - Scaling up the Ring AMCDs would call for a dual suspensor/drive station every 3 m along the device's circumference. The vertex station AMCDs assume a 1 cm magnetic gap, as do all EM systems throughout the fleet. But they develop an effective gap of ± 30 cm by using PZ linear actuators to fine-tune the relative position of the suspensor/drive electromagnets which make the gap itself. We will use this technique elsewhere in the fleet similarly to buy extra performance from the AMCDs. The power consumed by a PZ linear motor is unrelated to its stroke, but is related to its robustness. The AMCD rims we specify for the vertex station craft are six times heavier per unit length than those in the Ring craft, so their suspensor/drive stations must be comparably more robust, in addition to having stronger magnetics. We budget an undoubtedly conservative 9.5 kW per dual station.

10 MW

Propulsion - Stations 2,3,4 and 5 each have 1408 50 cm xenon ion engines. Each engine consumes 30 kW when operating at maximum thrust (performance data are taken from the references cited in Chapter 9). Allowing for about a fourth of the engines consuming their maximum simultaneously justifies our total power budget.

10 MW

Nervous System - Because of the unprecedented control system we posit in Chapter 8, no conventional accounting method for datonics power can apply defensibly to our fleet. Instead we take the unconventional but appropriate approach of following the biological paradigm, presuming that this at least leads to a conservative figure. We see in Chapter 8 that the human brain consumes about 25 W; the body gives off 85.6 W

sensible heat [Schubert et al, 85]. Now although there is more to the mammalian nervous system than the brain alone, our fleet controller is much less complex than a mammalian brain. So we will use this power ratio to budget our fleet intelligences at 30 % of the power total required to run the rest of the craft. In the particular case of the vertex stations, we budget 35 % since these resonator satellites need extra intercraft sensors and "extrinsic" intelligence to generate the laser beam at all, quite apart from controlling their own motions enough to do so.

69 MW

Finally, we assume that the subtotal of these power requirements represents what is left after 15 % system losses (due to transmission inefficiencies over the large dimensions of the satellites), and we increase the resulting before-losses figure with a growth margin of about 30 %.

$$\sum_{\text{power}} = \underline{405 \text{ MW}}$$

Mass Estimate

Power Plant - Our 405 MW rating is far enough out in the fringes of published parametric studies for space nuclear power systems (such as those referenced in Chapter 9) that system mass values which vary by well' over an order of magnitude can be justified bibliographically. Most such studies presume thermodynamic conversion plants in any case, with the system mass almost totally dominated by radiator mass. Our situation is quite different; as explained in Chapter 8, for reasons of reliability, maintainability, decentralization and vibration control we have chosen thermoelectric conversion, the least efficient. Our system mass will not be dominated by its radiators any more than are the small systems which make up its modules. The standard specific mass value of 35 kg/kWe for TE systems, based on 7 % efficiency,

leads directly to our budgeted value, since although our plant is extensive, it really is still just an assemblage of small plants.

14,000 MT

Mirror Segments - The 180 kg beryllium mirror segments we detail in Chapter 7, which are the "payload" of the resonator satellites, almost completely dominate their mass. The elliptical reflector area of the basic vertex stations required by their angled interception of the circulating beam, has area $\pi(500)(850) = 1.34(10^6) \text{ m}^2$. Each hexagonal segment has area 5.85 m^2 (Appendix A7-3); simple division indicates 228,235 mirror segments to cover the reflector area. Now there are small gaps between the segments, which reduces the number slightly, and some partial hexagons at the edges of the reflector, which increases the number slightly. In any case, we allow prudently for more than the bare minimum by specifying 230,000 segments.

41,400 MT

EM Isolation Mounts - There are 3 per segment, and we estimate 1 kg per actuator, based on the references in Chapter 8 but allowing for some material optimization.

690 MT

Mirror Support Structure - The 21 active members we have allocated for each mirror segment comprise 42.3 m of length. Appendix A4-7 details a mass breakdown for active truss members, which results in a specific mass of 0.67 kg/m, the value we use here also.

6,520 MT

AMCD Exoskeleton - As we did for the power rating of this active structure, we develop its mass by scaling up the bus mass of the Ring.

250 MT

Reactor Support Structure - We allow the probably excessive amount of 5 % of the mirror support structure mass.

330 MT

AMCD - We somewhat arbitrarily choose a rim mass 100 times greater than those of the Ring, which because of the much larger diameter means

that they have a cross-sectional area only about 6 times as great. Hence their sectional dimensions are roughly $2\frac{1}{2}$ times larger and thus do not particularly stretch the state of the art beyond the demonstrator AMCD rims referenced in Chapter 9. Both rims together mass 700 MT. The chase enclosure is a 3 m square section made of removable meteoroid bumper panels having a 1 cm total thickness of foamed reactive metal, surfaced by a thermal control finish. The enclosure mass is $2\pi(850)4(3)(.01)135 = 87$ MT. There are 1780 suspensor/drive stations in the chase, spaced at 3 m intervals. For each we budget 70 kg for the electromagnet assemblies and PZ mounts, 70 kg for fittings and structural attachments, including those which mount the enclosure panels, and 15 kg for various sensors.

1,100 MT

Resonator Sensors - These include all the non-structural sensors which the fleet controller needs to operate the planetary resonator. As outlined in Chapter 8, a hierarchy of optical, and particularly interferometric, sensors measures displacements between adjacent mirror segments, families, groups, neighborhoods, and regions, and monitors the reflector reference plane's location with respect to the cavity field. Some of these sensors are mounted on light, actively straight masts which project above the reflector plane; many are distributed throughout the mirror support structure behind the segments. We budget 5 kg per segment for these various sensors.

1,150 MT

Propulsion - The literature referenced in Chapter 9 indicates that a mass per engine on the order of 200 kg allows generously for the high-power 50 cm ion engines themselves, their power harnesses and mounting hardware, their beam neutralizers and dry tankage with its associated plumbing. For 1408 engines this amounts to 280 MT. We add about 2,700 MT of xenon propellant.

3,000 MT

Nervous System - This includes not only the sensory, processing and motor nerve circuits to run the craft itself, but also those to control the intracavity beam sensors, those necessary for exchanging intelligence with the other resonator craft and other members of the fleet, and the power distribution lines. Typically in the fleet we budget a mass equal to 20 or 25 % of the active structure mass for the nervous system, but for the vertex stations we budget 35 %. This allows for those extra laser-specific sensors just mentioned, as well as the mass penalty incurred by distributing power among and from such a widely dispersed reactor plant.

2,500 MT

Finally we inflate the mass subtotal with a 34 % growth margin, to reflect the general level of uncertainty with which we have outlined such an advanced space system.

$$\sum_{\text{mass}} = \underline{95,000 \text{ MT}}$$

Mass Properties

The vertex station is radially symmetrical. We establish a cartesian coordinate system having x axis through the geometrical center of and normal to the reflector surface. The y axis is parallel to the ellipse major axis, and the z axis makes a right-handed triad with the other two. The spacecraft centroid lies along the x axis, from symmetry. The ship is configured such that the power plant, propulsion system, AMCDs and supporting hardware all lie symmetrically disposed about the y - z plane, which passes through the centroid. To first order the sensors are also taken to be symmetrically located, so that only the mirrors, mirror support structure and their associated systems determine the location of the y - z plane. The mirror segments are 0.25 m thick, the structure is 4.2 m thick (but because its layer closest to the mirrors is denser, we take its own centroid to lie

1.8 m from the close face), and we allow a 0.1 m gap between them. Then the centroid lies:

$$\frac{41.4(.125) + (.69)(.2505) + (6.52 + 2.5)(.25 + .01 + 1.8)}{41.4 + 0.69 + 6.52 + 2.5}$$

$$= 0.47 \text{ m behind the front face of the mirror surface,}$$

or embedded 0.21 m into the mirror support truss. This metric defines the plane which bisects all other spacecraft systems in our reference configuration.

Because the y and z dimensions of the ship exceed the x dimension by over two orders of magnitude, in estimating inertial properties we consider the spacecraft as a flat plate of no thickness. The resulting errors for I_{yy} and I_{zz} are less than 1 %. Calculating the moments of inertia is simplified by considering the power plant and its associated systems as evenly distributed in the area between the reflector ellipse and the AMCD ring. We first derive simple algebraic expressions for the moments of inertia of that geometrical shape, by subtracting an ellipse from a disk:

$$I_{xx,\text{residual}} = \frac{(\pi r^2 m) r^2}{2} - \frac{(\pi a b m)(a^2 + b^2)}{4}$$

where $r \equiv$ disk radius, $a \equiv$ ellipse semimajor axis, $b \equiv$ ellipse semiminor axis, and $m \equiv$ homogeneous mass per unit area. Substituting the definition of m :

$$m \equiv \frac{M}{\pi r^2 - \pi a b}$$

where $M \equiv$ the (in our case known) total mass of that residual shape, recognizing that for us $a = r$, and simplifying yields directly:

$$I_{xx,\text{residual}} = \frac{M}{4} \left(\frac{2r^3 - b(a^2 + b^2)}{r - b} \right)$$

An analogous derivation produces expressions for the other two moments of inertia of the residual shape:

$$I_{yy, \text{ residual}} = \frac{M}{4} \left(\frac{r^3 - b^3}{r - b} \right)$$

$$I_{zz, \text{ residual}} = \frac{M}{4} \left(\frac{r^3 - a^2 b}{r - b} \right)$$

Now we have everything necessary for the inertial calculations. We lump the entire sensor system mass and nervous system mass budgets together with the mirror systems in the central ellipse, and in addition assume the propulsion system mass to be located at, rather than slightly beyond, the AMCD radius. We multiply the results by a factor of 1.34 to reflect the 34 % mass growth margin listed earlier.

$$\begin{aligned} I_{xx} &\approx \left[\frac{(41.4 + 6.52 + .69 + 1.15 + 2.5)}{4} (850^2 + 500^2) + \right. \\ &\quad \left. (.25 + 1.1 + 3.0) 850^2 + \right. \\ &\quad \left. (14 + .33) \frac{1}{4} \left(\frac{2(850)^3 - 500(850^2 + 500^2)}{850 - 500} \right) \right] (10^6)(1.34) \\ &= 31.4 \text{ T kg m}^2 \end{aligned}$$

$$\begin{aligned} I_{yy} &\approx \left[\frac{(41.4 + 6.52 + .69 + 1.15 + 2.5)}{4} (500^2) + \right. \\ &\quad \left. \frac{(.25 + 1.1 + 1.5)}{2} 850^2 + \right. \end{aligned}$$

$$+ \left(\frac{14 + .33}{4} \right) \left(\frac{850^3 - 500^3}{850 - 500} \right) (10^6)(1.34)$$

$$= 12.5 \text{ T kg m}^2$$

$$I_{zz} \approx \left[\left(\frac{41.4 + 6.52 + .69 + 1.15 + 2.5}{4} \right) (850^2) + \right. \\ \left. \left(\frac{.25 + 1.1 + 1.5}{2} \right) 850^2 + \right. \\ \left. \left(\frac{14 + .33}{4} \right) \left(\frac{850^3 - 850^2(500)}{850 - 500} \right) \right] (10^6)(1.34)$$

$$= 17.5 \text{ T kg m}^2$$

where the third term in each bracketed set incorporates the expressions worked out earlier for the residual area between the outer circle and its inscribed ellipse.

Appendix A4-2 Stations 2, 3, 4 and 5 system performance.

AMCD Authority

The maximum working spin rate allowed the Kevlar 850 m radius rims is 1.45 rad/s (Appendix A9-4). Because rim speed changes produce roll maneuvers, we avoid starting the mission in a roll-saturated condition by biasing the rims 10 % below their maximum speed. Each rim therefore stores angular momentum:

$$H = I\omega = mr^2\omega = 350,000(850)^2(.9)1.45 = 330 \text{ G N m s}$$

We have specified ± 30 cm effective gap tilt for these rims, which means that with the counter-rotating rims tipped oppositely to the limit of their stroke, the available angular momentum for precessional torquing is:

$$H = 2(330)(10^9) \sin \tan^{-1} \left[\frac{.30}{850} \right] = 233 \text{ M N m s}$$

This can turn the craft about any axis normal to the x axis. The slowest turn (worst-case) would be about the z axis:

$$\omega = \frac{H}{I_{zz}} = \frac{233(10^6)}{17.5(10^{12})} = 13.3 \text{ } \mu\text{rad/s} = 1.27(10^{-4}) \text{ rpm}$$

At this rate, a quarter turn would take about 33 hr. Because of its planet-oriented attitude, the craft must of course actually make a full turn every two hours, but that motion represents the bias free-body rotation with which the ship begins its mission life, imparted to it by external means at the time of construction. The rotation authority we have just calculated is for trimming that bias, and at 1.5 % of that bias, we expect it to be sufficient. The principal-axis orientation of the basic vertex stations means that gravity-gradient

torques will balance, thus aiding stabilization.

Rolling about the x axis (a yawing maneuver in the orbital frame) is accomplished by differential rim acceleration. If we require this rate to match the rate just calculated, it would take angular momentum:

$$H = I_{xx} \omega = 31.4(10^{12}) 13.3(10^{-6}) = 418 \text{ MN ms}$$

effected by a differential rim speed of:

$$\omega = \frac{H}{mr^2} = \frac{4.18(10^8)}{(350,000) 850^2} = 1.65 \text{ mrad/s}$$

representing just a:

$$\frac{\left(\frac{1.65(10^{-3})}{2} \right)}{1.45} 100 = 0.057 \% \text{ speed change for both rims.}$$

The power implications of rim speed changes are investigated in Appendix A4-4.

Propulsive Capacity

Appendix A8-5 shows that exospheric drag yields a total force measured in less than mN, even for the largely frontal satellites like 1β. In addition, using the value for the solar wind force from Chapter 8, we calculate a pessimistic but nonetheless trivial maximum total force of $4.4(10^{-9})\pi 850^2 = 10 \text{ mN}$ on the vertex stations. Both of these are utterly dominated by the forces of light pressure and solar gravity.

The solar tug is always directed toward the sun, while the net pressure from radiation (Appendix A8-6) is directed away from the sun at worst. Since these two forces oppose each other on the dayside, the most severe force requiring instantaneous compensation is the solar tug alone during darkside passage. Using the value from Appendix A8-3,

we calculate the peak force to be $1.6(10^{-6})95(10^6) = 152 \text{ N}$.

Our 50 cm ion engines are rated for a maximum power of 30 kW, with efficiency of 0.8 and specific impulse of 4500 s. Thus each can produce maximum thrust:

$$T = \frac{2\eta P}{I_{sp}g_o} = \frac{2(.8)30(10^3)}{4500(9.8)} = 1.1 \text{ N}$$

At least 140 engines would be needed to get 152 N. We specify that this number represent 80 % of the ship's unidirectional capacity so that a total of 176 such engines are available for the job. By dividing this complement among two locations (for redundancy and to avoid unwanted propulsive torques) we arrive at 88 engines per direction for each of the four engine outriggers. The total number is thus 1408 engines (this quad-configuration actually doubles the number available for thrusting normal to the reflector face).

To analyze the propellant stores longevity, we need to know the average "constant" force which the engines must provide. Following the procedure outlined in Appendix A8-3 and allowing for the slight dayside counterthrust of radiation pressure, we find an average force of:

$$\frac{152 + (152 - 10)}{\pi} = 94 \text{ N}$$

Now the thrust T of an engine is simply:

$$T = u\dot{m}$$

where $u \equiv$ the exhaust velocity and $\dot{m} \equiv$ the mass flow rate. The exhaust velocity in turn is found from the specific impulse rating:

$$I_{sp} = \frac{u}{g_o}$$

We combine these relations to find the total propellant mass which must

be expelled to result in an effectively continuous 94 N:

$$\dot{m} = \frac{F}{I_{sp}g_o} = \frac{94}{4500(9.8)} = 2.12 \text{ g/s} = 67,220 \text{ kg/yr}$$

$$= 672 \text{ MT} / 10 \text{ yr}$$

If we add enough additional propellant to allow for a general station-keeping budget of $\Delta v = 100 \text{ m/s/yr}$ for those same 10 yr, we need another:

$$m_p = m_i \left[1 - \exp\left(\frac{-\Delta v n}{I_{sp}g_o}\right) \right] = 95(10^6) \left[1 - \exp\left(\frac{-100(10)}{4500(9.8)}\right) \right]$$

$$= 2130 \text{ MT}$$

The sum of these two 10 yr propellant stores and the propellant system hardware itself constitutes our 3,000 MT system mass.

Mirror Actuation

The EM mirror segment isolation mounts, in addition to performing the mode-hopping focus changes necessary to allow continuous lasing, also tilt the mirror segments to compensate through micro-aiming for gross resonator satellite relative displacements due to planetary gravity variations. If we assume two consecutive satellites undergo simultaneous worst-case displacements as estimated in Appendix A8-1, we can follow the procedure of Appendix A7-12 to find the actuator stroke necessary to compensate:

$$\frac{710(\sin 36^\circ) + 2200(\cos 36^\circ)}{8983(10^3)} \frac{2.1}{2} = 257 \text{ } \mu\text{m} = 0.26 \text{ mm}$$

easily provided by the powerful EM actuators we have specified.

Appendix A4-3 1δ & 1ε systems sketch.

1ε Power Requirement

Systems and their power requirements for this satellite are closely based on those of the Ring (Appendix A4-7).

Propulsion - 1ε has a full complement of 24 20-cm xenon ion engines, each rated at 5 kW for 0.2 N maximum thrust. Assume 8 engines operating simultaneously at their maximum.

40 kW

Bus Motor Control - The active bus structure is identical with that of the Ring.

102 kW

AMCD and Mirror Control - The AMCDs are identical to those in the Ring, so their total power requirement is also 40 kW. To this we add 10 kW for the two mirror mounting assemblies. Each consists of both a linear EM-driven turntable pivot (for fast mirror rotation) and an EM final stage mirror isolation space bearing (for fine-pointing).

50 kW

Nervous System - Following the biological analogy of Appendix A4-1 for sensor and processing network power, we budget 30 % of the power subtotal.

58 kW

Assuming power system losses consume 10 % of the available electrical power and that a 20 % growth margin is appropriate for this satellite leads us to inflate the total.

$$\sum_{\text{power}} = \underline{333 \text{ kW}}$$

1δ Power Requirement

AMCD and Tether Bearing - The AMCDs are identical to those in the Transducer (Appendix A4-5). Although we account for mirror actuation power elsewhere, here we budget about 4 kW for the crawler motors and EM space bearing positioning actuators in the tether attachment assembly.

8 kW

Propulsion - Although 1δ ends up massing three times as much as the Transducer, its station-keeping needs are less severe since mobility is unnecessary. Thus we budget for 4 0.1 N 20-cm engines operating simultaneously.

10 kW

Mirror Control - 1δ's reflective surface is an off-axis paraboloid made up of 1 m hexagonal segments. Each such segment has area 0.65 m^2 , calculated as shown in Appendix A7-3. Covering an area of radius 7.5 m with these takes about 270 of them. Assume an EM mirror isolation mount consumes an average of 12 W. Given three mounts per mirror, the total is 9.7 kW.

10 kW

Mirror Support and Bus Motor Control - The bus armature connecting the two ends of the satellite and encircling the tether bearing consists of about 16 bays, each of which can be considered to have 14 members. Allow a total of 250 members, each of which consumes 20 W under active control (Appendix A4-7) for a power budget of 5 kW. The mirror support structure is an equally active, redundant assembly several layers thick. We allocate 25 members through the support thickness for each mirror segment, for a total of 6750. Since these members are in general much shorter than the bus members, we allow 10 W for each; the power budget is 68 kW. We increase the 73 kW subtotal by a third to account for the power consumption of the active nodes which

join the members (Appendix A4-7).

97 kW

Nervous System - As usual we budget an additional 30 % of the power subtotal.

38 kW

We allow 10 % system losses, and inflate the total by a 20 % growth margin.

$$\sum_{\text{power}} = \underline{217 \text{ kW}}$$

1e Mass Estimate

Because the bus structure, AMCDs and distributed controller are identical to those of the Ring, we adopt their values intact from Appendix A4-7.

Bus Structure - 8000 kg

AMCD - 14,400 kg

Nervous System - 1600 kg

Propulsion - The 24 ion engines together mass 600 kg, which we take to include the engines themselves, their beam neutralizers, power harnesses and mounting hardware, as well as the outriggers for those clusters accomplishing ring roll and in-plane translation. The dry tankage mass will exceed that for the Ring, since we call for three times as much tanked xenon (we expect the more severe near-Venus torque and drag perturbations to require more propulsive effort).

10,000 kg

Mirror Assemblies - Each of the identical assemblies consists of an elliptical, flat monolithic cored mirror, attitude, pointing and beam sensors, EM space-bearing isolation mounts, mounting strongback with its strain sensors and actuators, mission-associated datonics, and the

power- and signal-transmitting EM-driven turntable pivot mounting the assembly to the bus structure.

8000 kg

Because of the inherent uncertainty in such a brief sketch of an advanced space system, we inflate the total by a 31 % growth margin.

$$\sum_{\text{mass}} = \underline{55,000 \text{ kg}}$$

1δ Mass Estimate

Power Plant - The out-of-core thermionic reactor supplies power to both the 1δ and 1ε satellites, transmitting it between them through the conductive tethers. Summing the power requirements for both satellites yields a 550 kWe power plant rating. The system mass for reactors in this class is about equally distributed among core, shields, power converters and radiator, and the total varies by up to a factor of 3, depending on whose data are used. Because of our emphasis on longevity, maintainability and thorough shielding, we choose a system mass toward the high end of the range.

18,000 kg

Mirrors - Segment mass is taken directly from the demonstrated values quoted by the references of Chapter 7: 20 kg.

5400 kg

Mirror Support and Bus Structure - If we assume an average member length within the mirror support truss of 1 m, and an average member length within the bus truss of 2 m, then the total structural material length can be taken as $(270)(25)1 + (16)(14)2 = 7200 \text{ m}$. The specific structural mass derived in Appendix A4-7 for active truss equipment (including composite tubular members and nodes, thermal actuators, PZ multimorph coatings, service fittings, and fiberoptic strain and temperature sensors) amounts to 0.67 kg/m of member length. We use

this value here, although since the mirror support structural elements are more gracile than those of the Ring exoskeleton, it probably represents an overestimate.

5000 kg

EM Mounts and Tether Bearing - Allotting 2 kg for one isolation and fine-pointing fixture yields a total of $(270)(3)2 = 1620$ kg. The tether bearing includes conducting gripper wheels, crawler motors and circuitry, the bearing framework, power transfer cables, and the tunable EM gimbal mount; we allocate another metric ton.

2600 kg

AMCD - The suspensor/drive stations are essentially the same as those for the Transducer (Appendix A4-5), but we make the rims twice as massive (1000 kg each) to gain greater control authority for this more massive spacecraft. We specify foamed reactive metal debris bumpers (Appendix A4-7) to define the rectangular section chase, adding a mass of $(135)\pi 15(2+2+1+1)(.01) = 382$ kg. We add allowance for dedicated optical sensors in the chase and other specialized fittings, as well as the structural mounts for the figure sensor masts.

3000 kg

Propulsion - 1δ has 24 20-cm xenon ion engines for redundant and fully 6-DOF propulsion. At 25 kg/engine for all associated plumbing and hardware, the total is 600 kg. To this we add 2000 kg for tanked propellant (much less than for 1ε, even though their masses are comparable, since 1δ has tether-actuation available and can thus offload some of its propulsive burden to 1ε).

2600 kg

Nervous System - Targeting sensors and mission datonics, bus inertial sensors, power distribution lines, and sensory, processing and motor nerves comprise a total mass assumed equivalent to 25 % of the active structural mass.

1250 kg

Finally we inflate the mass total with a 32 % growth margin.

$$\sum_{\text{mass}} = \underline{50,000 \text{ kg}}$$

Tether Mass Estimate

As discussed in Chapter 9, tether mass is dominated by its power conducting function rather than by its structural function. For reference we will use the 922 kg/km specific mass found in the literature cited in Chapter 9 for a primitive aluminum tether rated at 500 kW. A 2 km length would mass 1844 kg. Unlike a typical electrodynamic tether, however, ours does not use ionospheric coupling to complete its circuit, but requires instead a separate return path. Twice 1844 is 3688 kg. Because we anticipate using a more advanced tether construction (as discussed in Chapter 9) we reduce this value.

3000 kg

1E Inertial Properties

Choose a cartesian coordinate system with x axis concentric with the bus, z axis parallel to the diameter connecting the two mirror assemblies, and y axis forming a right-handed triad with those two. To find the origin (centroid), first note that it must lie along the x axis because the spacecraft is reflectively symmetric about that axis. The y - z plane, however, will not bisect the ring plane since both mirror assemblies lie on the same face of the ring. The centroid itself is:

$$\frac{8(10)}{8 + 14.4 + 1.6 + 10 + 8} = 1.9 \text{ m}$$

away from the bus central plane on that same side.

Since during nominal operation the 1c ring plane must lie parallel to the resonator orbit plane itself, the tether attach-point, representing as it does the site of a constant force on the bus, must also lie in the $y - z$ plane. That is, the line connecting the centroid with the point at which the tether tension force is applied must at all times be maintained parallel to the bus plane, or the AMCDs will become saturated and the xenon reserves depleted. The tether attachment mechanism, by translating under active control, adjusts the actual attach-point to compensate for calibration bias and CM-shift during the mission life.

We model the bus as a thin ring containing all the distributed mass of the spacecraft save the two mirror assemblies and the four engine clusters; those systems are considered as point masses located at their respective centroids. Then:

$$I_{xx} \approx (55,000) 50^2 = 138 \text{ M kg m}^2$$

$$\begin{aligned} I_{yy} &\approx (1.31) \left[\left(\frac{1}{2}(8000 + 14,400 + 1600) + 8000 \right) 50^2 + \right. \\ &\quad \left. (10,000) 35.4^2 \right] \\ &= 82 \text{ M kg m}^2 \end{aligned}$$

$$\begin{aligned} I_{zz} &\approx (1.31) \left[\frac{1}{2}(8000 + 14,400 + 1600) 50^2 + (10,000) 35.4^2 \right] \\ &= 56 \text{ M kg m}^2 \end{aligned}$$

where contributions due to the parallel axis theorem have been ignored (the error in this case is about 0.15 %), and the factor 1.31 inflates the tabulated mass values as discussed previously.

16 Inertial Properties

Choose a cartesian coordinate system having x axis along the bus (symmetry) axis, y axis concentric with the tether bearing gimbal ring, and z axis forming a right-handed triad with those two. The centroid will of course lie along the x axis, and its location will determine the designed nominal location of the teather bearing. To find it, note from the mass analysis that 94 % of the structure mass is located at the bow (supporting the segmented mirror). Therefore ignore the remaining 6 % which is the bus strongback (its effect on the centroid location is inconsequential in this case). Similarly, 16 of the 24 engines, and thus $2/3$ of the propulsion system mass, is also located at the bow; the rest is located symmetrically about the centroid in any case. The centroid is about:

$$\frac{(30) 18}{18 + 5.4 + (.94)(5 + 1.25) + 1.6 + 3 + (.67)2.6} = 15 \text{ m}$$

aft of the bow, or at the bus midpoint. In the sequel we continue to approximate the bus strongback as massless, and model the other systems as disks, rings and point masses, as appropriate. Then:

$$\begin{aligned} I_{xx} &\approx (1.32) \left[\frac{1}{2} ((18,000)2^2 + (5400 + 5000 + 1250 + 1600)7.5^2) + \right. \\ &\quad \left. (3000 + (.67)2600)7.5^2 + (.33)(2600)3^2 \right] \\ &= 902 \text{ k kg m}^2 \end{aligned}$$

$$\begin{aligned} I_{yy} &\approx I_{zz} \approx (1.32) \left[(18,000)15^2 + \right. \\ &\quad \left. (5400 + 5000 + 1600 + 1250) \left(\frac{7.5^2}{4} + 15^2 \right) + \right. \end{aligned}$$

$$\begin{aligned}
& + 3000 \left[\frac{7.5^2}{2} + 15^2 \right] + (.67)2600(16.8^2) + \\
& \quad (.33)2600(3^2) \Bigg\} \\
& = 11.2 \text{ M kg m}^2
\end{aligned}$$

where parallel axis contributions dominate, and algebraic contributions from rigid-body rotations of the submasses have been ignored when they amount to less than 2 % of the dominant terms.

Tethered Formation Mass Balance

Optical constraints from Chapter 7 favor locating 1ε 1000 m above the resonator orbit, and 1δ 800 m below the orbit. Since 1δ is the lighter of the two and the least distant, we insure that the tethered formation flies with its CM on the resonator orbit by suspending a passive counter-mass below 1δ. We station the mass 200 m below 1δ, or 1000 m below the orbit. It must therefore have mass:

$$\frac{(55,000)10 - (50,000)8}{10} = 15,000 \text{ kg}$$

All kinds of uses for a 15 Mg planet-oriented satellite could be envisioned; for example, powered by RTGs, it could perform completely autonomous planetary science. However, for reference we will consider it to be completely inert. If made of lead, the counter-mass would be a sphere only 1.4 m across, essentially immune to exospheric drag.

Appendix A4-4 1δ & 1ε system performance.

Tether

The tether comprises less than 3 % of the paired system mass. If its own mass is ignored, then the tension can be approximated (for all but extremely long tethers) easily, and is equal to the gravity gradient force [Barakat & Butner, 86]:

$$T = F_{gg} \approx 3 L m \frac{\mu}{r_o^3}$$

where $L \equiv$ distance from system CG at which the mass m is attached, $\mu \equiv$ planetary gravitational parameter, and $r_o \equiv$ orbital radius of the system CG. Substituting our values and using 1ε:

$$T \approx 3(1000)(55,000) \frac{3.45(10^{14})}{(7641(10^3))^3} = 130 \text{ N}$$

Our conductive requirements result in a reference tether structure excessively safe by any standards for this modest load. For comparison we note that two 2 cm aluminum tethers have a cross-sectional area of $6.3(10^{-4}) \text{ m}^2$; since the longitudinal tensile strength of high-grade aluminum may be taken as 400 MPa, these could support 0.25 MN, or almost 2000 times as much as we require. Tether designers usually employ a safety factor of 3.5.

Libration damping and CG tuning both require effective tether length changes under active control. We accomplish this by having 1δ crawl up and down on the tethers. For example, a 1 m shift of 1δ yields a:

$$\frac{55 - (.801)50 - 15}{-(55 + 50 + 15)} 1000 = 0.42 \text{ m shift of the system CG.}$$

Since the constant 130 N tether tension force represents a potentially huge attitude perturbation source, it is interesting to compare its effect to the other constant force we have designed into the spacecraft, namely their ion engines. With four engines operating at maximum thrust to pivot 1ϵ about its x axis, for example, available torque is $4(.2)50 = 40$ Nm; the tether can match this torque merely by shifting its attach-point by $40/130 = 0.3$ m. A similar analysis for 1δ shows that the tether can match maximum engine torque by shifting only 5 cm! Clearly, active control of the attach-point is required to avoid a rapid buildup of secular attitude torques due to tether tension. However, that same control, when applied willfully, comprises a powerful attitude stabilization tool.

1 δ AMCD Authority

Critical performance occurs four times per orbit, when Station 1 crosses over the planetary terminators, subsolar point, and antisolar point. At these times the intermediate beam must be switched from one rim mirror on 1ϵ to the other (Figure 7-12), an angular distance of 56 mrad. Although the 15 m diameter Kevlar rims can spin as fast as 164 rad/s (Appendix A9-4), we bias them at only 90 % of that speed so they are not saturated as the mission begins. Each rim stores:

$$H = I\omega = mr^2\omega = 1000(7.5)^2(.9)164 = 8.3 \text{ MNms}$$

of angular momentum. The suspensor/drive stations use PZ magnet actuation to augment the 1 cm available magnetic gap to an effective gap of ± 3 cm. The maximum angular momentum available for precessional torquing if both rims are tipped fully is therefore:

$$H = 2(8.3)(10^6) \sin \tan^{-1} \left(\frac{.03}{7.5} \right) = 66.4 \text{ kNms}$$

Tipping the rims in this way can turn the spacecraft about any axis normal to its x axis at a speed of:

$$\omega = \frac{H}{I_{yy}} = \frac{66.4(10^3)}{11.2(10^6)} = 6.0 \text{ mrad/s} = .06 \text{ rpm}$$

or greater. Thus the pivoting maneuver to retarget the intermediate beam to 1ϵ 's opposite rim mirror (a maneuver requiring the 1δ reflector and therefore the spacecraft to rotate an angular distance equal to half the 56 mrad target separation, or 28 mrad) can be accomplished in only $28/6 = 4.75$ s. Of course, the AMCD rims cannot be tipped instantaneously, so the acceleration and deceleration times limited by the PZ actuators in the suspensor/drive mounts will extend the practical slewing time. We specify 20 s for retargeting.

1 ϵ AMCD Authority

Following exactly the algebraic procedure outlined in the last section, but using 100 m diameter rims which mass 3500 kg each and are limited to speeds below 12.3 rad/s (Appendix A9-4), we find that each rim in 1ϵ stores 97 MNms of angular momentum. With 3 cm effective gap tipping, 116 kNms is available for precessional torquing, which can tilt the 1ϵ bus ring at speeds up to 1.4 mrad/s, or 0.014 rpm. However, it is the in-plane rotations which are of particular concern. Over the planetary terminators (points A and C in Figure 7-12), 1ϵ switches the beam between L1 and L2 Stations. So as 1δ retargets the beam across the diameter of 1ϵ , 1ϵ must rotate to insure that its rim mirrors will be "out of their own way" during the next quarter orbit. In preparing for that interval of unobstructed viewing, 1ϵ must rotate through twice the angle subtended by the 10 m intermediate beam diameter at the distance of the opposite rim mirror, or 200 mrad. If this "roll" maneuver is to take place during the same 20 s that 1δ takes to retarget the beam, thus minimizing downtime,

then it must occur at an average speed of $200/20 = 10$ mrad/s.
 Rotating 1ε at this speed takes angular momentum:

$$H = I_{xx} \omega = 138(10^6)10(10^{-3}) = 1.4 \text{ MNms}$$

which is extracted from the momentum reserve stored in the counter-rotating rims by accelerating one and braking the other. The differential speed must be:

$$\Delta\omega = \frac{H}{mr^2} = \frac{1.4(10^6)}{3500 (50)^2} = 160 \text{ mrad/s}$$

which represents a 0.71 % speed change for both rims. One of the best features of an AMCD is that its EM drive facilitates rapid speed changes. To gain an appreciation of this, first note that the energy required to accelerate one of the rims to 1.0071 its nominal speed is

$$\begin{aligned} \Delta KE &= \frac{1}{2} I(\omega_2^2 - \omega_1^2) \\ &= \frac{1}{2}(3500)50^2[(.9)(12.3)]^2(1.0071^2 - 1) = 7.6 \text{ MJ} \end{aligned}$$

which is really an enormous amount. For instance, if it were to be delivered in 2 s by an external source, the power requirement would be 3.8 MW, or 7 times as much as the reactor which supplies both 1δ and 1ε could provide! Fortunately, a dual-rim AMCD is a conservative system; energy extracted by braking one rim is used to accelerate the other. Thus except to make up that energy lost through system inefficiencies (mostly joule heating and magnetic drag), no external power is required to execute a major roll maneuver. Detailed study is necessary to characterize the real effect accurately; even a 1 % conversion inefficiency would consume 38 kW for those 2 s, most of the AMCD power budget.

Propulsive Performance

We noted in Appendix A4-3 that the tethered link connecting 1 δ and 1 ϵ allows us to combine their propulsive burdens. The perturbation appendices of Chapter 8 show that third-body gravitational effects constitute the worst secular disturbance force which the tethered pair must compensate propulsively. Specifically, we see from Appendix A8-3 that the acceleration resulting from the solar gravitational potential amounts to a constant 1.0 $\mu\text{N/kg}$, and the acceleration due to the perturbing gravitational presence of 1 γ nearby amounts to about 37 $\mu\text{N/kg}$. Thus the total propulsive force required from the tethered pair to counteract these tugs is:

$$(55 + 50)(10^3)(1 + 37)(10^{-6}) = 4 \text{ N}$$

applied constantly. The mass flow rate required to provide such a force, assuming a specific impulse of 4100 s, is then:

$$\dot{m} = \frac{F}{I_{sp}g_o} = \frac{4}{4100(9.8)} = 1.0(10^{-4}) \text{ kg/s} = 3140 \text{ kg/yr}$$

This rate will consume our total propellant budget of about 10,500 kg in a little over 3 years. The two options available, whose resolution we leave unstudied, are either to increase the propellant stores in the next systems design iteration, or to reconfigure the intercraft architecture to incorporate 1 δ and 1 ϵ into the dispersed structure which we use to keep 1 α , 1 β , and 1 γ apart.

Appendix A4-5 Transducer systems sketch.

Power Requirement

Modulator - Assume by analogy with transistor gain that the beam can be modulated using about 1 % as much power as it contains. This budget is $(.01)(180) = 1.8$ kW. Inflate it to account for membrane strain sensors and frame sensors and strain actuators.

2 kW

Propulsion - Ion engine power consumption varies with design specific impulse. Adapting data from the sources referenced in Chapter 9, assume 20 cm xenon engines, each using 2.5 kW maximum to produce a maximum thrust of 0.1 N. Station exchange with the Ring poses the most severe scenario (4 engines simultaneously, maximum thrust).

10 kW

Datronics - This includes mission processing, spacecraft control processing, state sensors (inertial devices and star trackers), and interstation C³I (optical links and beam pattern sensors).

5 kW

Control - The demonstration AMCD referenced in Chapter 8 uses suspensor drive stations at 2 m intervals along the rim, with three drawing a maximum of 4 A from a 60 V supply. Power per station is thus $(60)(4/3) = 80$ W. Each of our two 15 m diameter rims would have 24 similar stations, for a total maximum consumption of $2(24)(80)$ or 3840 W. To this we must add a budget for the PZ suspensor-drive mounts and the EM diaphragm frame fine-pointing mounts.

5 kW

Although superposition is conservative (the modulator is not used during maximum engine-thrust station-exchange maneuvers), the high

level of uncertainty and low level of detail in this sketch lead us not only to add these values but inflate the total with a growth margin as well.

$$\sum_{\text{power}} = \underline{30 \text{ kW}}$$

Mass Estimate

Power System - These values derive from the sources referenced in Chapter 9. Power system mass includes the reactor core, shield, power conversion hardware, thermal radiator, and power conditioning and distribution equipment. Typical specific power (including conversion inefficiency) may be taken as 40 - 55 W/kg. System mass for reactors smaller than about 100 kWe is approximately equally dominated by the core and the shield. Since our in-core thermionic converter is comparatively efficient, but our shield must be 4π and thicker than a conventional design, we take the conservative mass, and then double it: $2(30,000)/40 = 1500 \text{ kg}$

1500 kg

Modulator - This subsystem includes the diaphragm mirror membrane, its mounting frame, the transducer assembly between them, and the embedded sensors and strain actuators in all three. We estimate the mass of the diaphragm and its associated equipment by assuming a disk 1 mm thick of beryllium: $\rho V = 1.85(10^3)(10^{-3})\pi(7.5)^2 = 327 \text{ kg}$. Assume the beryllium frame with its sensors and actuators is five times as massive or 1635 kg. Budget one tenth of the subtotal for the transducer, 196 kg.

2200 kg

Propulsion - Electric engines are referenced in Chapter 9. Given that our configuration requires unusually long propellant feed lines from the tankage to the thruster modules, a reasonable per engine mass for

our 20 cm xenon ion engines, including an allocation for power harness, beam neutralizer, mounting hardware, feedlines, empty tankage and fittings, is 25 kg. The Transducer uses four groups of four engines each: $16(25) = 400$ kg. In addition, we budget another 600 kg for propellant.

1000 kg

Datonics - Being distributed throughout the spacecraft and surface-mounted on the back of the modulator, the mission and control datonics does not require a thermal rejection mass budget. Furthermore, its primarily optical componentry is presumed to be substantially lighter than comparable all-electronic circuitry.

600 kg

Sensors - The sensor complement not already included with the modulator consists of inertial accelerometers, fiberoptic strain and temperature sensors distributed throughout the bus, optical intercraft and star trackers, and the beam overspill collar with its active spar structure.

600 kg

Control - Assume the four EM fine-pointing mounts together mass 100 kg. We choose an AMCD rim mass/length almost $2\frac{1}{2}$ times greater than the Ball demonstrator's 4.5 kg/m (although we did not increase the power budget, to allow for reasonable system optimization maturity), for a rim mass of 500 kg. Each of the two rims has 24 suspensor-drive stations which, augmented from the Ball units by PZ positioning mounts, total 300 kg for the system. We budget another 600 kg for the debris-bumper channel enclosure panels.

2000 kg

Bus Structure - Assuming an average thickness of 0.1 m, the volume of material in the bus armature can be estimated by summing contributions from the reactor shell hub, in-plane spokes, out-of-plane spokes, and AMCD channels:

$$(\pi(1.2)^2 + 4(1)(7.5) + (4/3)\pi(7.5)^2 + \pi(15)^2)(.1) = 15 \text{ m}^3$$

From Chapter 9, a conservative density for foamed reactive metal is 135 kg/m^3 , so the armature mass would be $135(15) = 2050 \text{ kg}$. To this we must add an allowance for fittings and attachments, and the thermal and PZ actuator layers distributed across the armature surface.

3000 kg

The mass subtotal for these subsystems is 10,900 kg, which we will inflate by roughly the industry-standard 40 % to allow for mass growth of a preliminary, incompletely defined, new space system.

$$\sum_{\text{mass}} = \underline{15,000 \text{ kg}}$$

Inertial Properties

We choose a cartesian coordinate system with y and z axes parallel to the bus armature spokes, x axis along the spacecraft symmetry axis and origin in the inertial core. Due to radial symmetry, the centroid lies along the x axis, and we configure the reactor such that the centroid coincides with the coordinate origin, by balancing the mass moments of the power system and modulator. Thus the power system centroid must be about:

$$\frac{2200(1)}{(1500 + 500)} = 1.1 \text{ m aft of the inertial core.}$$

We estimate the spacecraft mass moments of inertia by approximating its mass elements as a set of rings, disks, and point masses. We overestimate by combining the sensor and datonics mass with the AMCD at the maximum radius, and locating the entire propulsion system mass at that radius as well:

$$\begin{aligned} I_{xx} &\approx \frac{1}{2}(2100)(.5)^2 + (\frac{1}{2}(3080) + 1400 + 4480 + \frac{1}{2}(4200))(7.5)^2 \\ &= 540,000 \text{ kg m}^2 \end{aligned}$$

where the masses of the power system, modulator, propulsion system, AMCD and structure respectively have been inflated 40 % as noted above. Similarly:

$$\begin{aligned}
 I_{yy} = I_{zz} &\approx 2100 \left[\frac{(.5)^2}{4} + \frac{(2.5)^2}{12} + (1.8)^2 \right] + \\
 &3080 \left[\frac{(7.5)^2}{4} + \frac{(1)^2}{12} + (1)^2 \right] + \\
 &\frac{1400}{2} (7.5)^2 + \\
 &4480 \left[\frac{(7.5)^2}{2} + \frac{(.5)^2}{12} + (.25)^2 \right] + \\
 &4200 \left[\frac{(7.5)^2}{4} + \frac{(.5)^2}{12} + (.2)^2 \right] \\
 &= 280,000 \text{ kg m}^2
 \end{aligned}$$

where the third term in each bracketed set derives from the parallel axis theorem.

Appendix A4-6 Transducer systems performance.

Three general categories of performance are of interest: overall attitude control authority, fine-pointing, and station-keeping. Together these establish the operating envelope within which the Transducer can do its job.

AMCD Authority

As derived in Chapter 9, the maximum spin rate for the 7.5 m radius Kevlar AMCD rim is 164 rad/s. To allow for "roll" saturation, bias the rims below their maximum, at 150 rad/s. Then each rim stores angular momentum:

$$H = I\omega = mr^2\omega = 500(7.5)^2 150 = 4.2 \text{ MN m s}$$

By a combination of electromagnetic gap torquing (for ultrafine) and physical gap relocation (PZ actuated for fine control), we specify a maximum controlled out-of-plane rim stroke for this system of 1 cm in either direction from the nominal. With both rims tipped maximally, the angular momentum available for precessional torquing is:

$$H = 2(4.2)(10^6) \sin \tan^{-1} \left(\frac{.01}{7.5} \right) = 11.2 \text{ kN m s}$$

The maximum angular control rate about the y (pitch) and z (yaw) axes is therefore:

$$\omega = \frac{H}{I_{yy}} = \frac{11,200}{280,000} = .04 \text{ rad/s} = .38 \text{ rpm}$$

The maximum rate about the x (roll) axis is of course unrelated to this value, being limited instead by the AMCD control acceleration and

the ability of the spacecraft active structure to withstand large impulsive torques. However, the angular momentum required to roll the spacecraft at the pitch and yaw rates just derived is:

$$H = I_{xx} \omega = 540,000 (.04) = 21.6 \text{ kN m s}$$

effected by a differential rim speed of:

$$\omega = \frac{H}{mr^2} = \frac{21,600}{500(7.5)^2} = .77 \text{ rad/s}$$

which represents just a 0.26 % speed change for both rims.

Pointing Dither

During the one continuous hour (half of the resonator orbit) that each Transducer is used, it must cancel the beam's 15 mrad angular dither caused by the resonator coupler crossing from one side of Venus to the other. If this were accomplished by the modulator EM mounts alone, they would need at least a:

$$\tan(15 \text{ mrad}) (7.5) = .11 \text{ m}$$

bidirectional stroke. Since the EM mounts are intended for fine-tuning this dither requirement would lead to excessively heavy and power-consuming mounts. Instead the bus AMCD will compensate about 99 % of the dither. The fastest compensation necessary occurs when the coupler cross-track velocity is greatest, at noon for L1 and midnight for L2. That maximum rate is then:

$$\frac{6.52 \text{ km/s}}{1(10^6) \text{ km}} = 6.5 \text{ } \mu\text{rad/s} \quad \text{which is over 6000 times smaller}$$

than the 40 mrad/s control authority we have available.

Station Keeping

Xenon ion engines operating at the power levels we have specified are capable of a specific impulse around 3500 s. Station-keeping requirements at an unstable libration point are directly related to the precision with which the spacecraft is placed at the point, but would in general not exceed a few m/s per year of Δv . Choosing an excessive Δv specification of 100 m/s per year, a ten year supply of propellant would amount to:

$$\begin{aligned} m_p &= m_i \left(1 - \exp \left(- \frac{\Delta v}{I_{sp} g_0} \right) \right) = 15,000 \left(1 - \exp \left(- \frac{10(100)}{3500(9.8)} \right) \right) \\ &= 431 \text{ kg} \end{aligned}$$

We have already budgeted 600 kg for propellant stores. The extra allotment permits occasional slight station adjustments to keep the Transducer itself out of the way of the Ring's redirected beam for particular target stars, and the station-exchange maneuvers which can be expected a few times a year. With four thrusters at maximum power, the Transducer will experience an acceleration of:

$$a = \frac{F}{m} = \frac{(.4)}{15,000} = 2.7(10^{-5}) \text{ m/s}^2$$

enough to move it 50 m in just:

$$t = \left(\frac{2(50)}{2.7(10^{-5})} \right)^{\frac{1}{2}} = 1925 \text{ s} = 32 \text{ min}$$

After an equal deceleration interval, the Transducer will have passed through the Ring and sufficiently beyond it to operate alone.

Appendix A4-7 Ring systems sketch.

Power Requirement

Propulsion - The Ring is a pseudo-passive reflector satellite (its secondary mirror only retargets the outgoing encoded beam), requiring simply stationkeeping propulsion. Thus despite its size it uses the same 20 cm xenon ion engines as the Transducer, albeit rated for a 0.2 N maximum thrust each, drawing 5 kW of input power. Running 8 engines simultaneously (for example during recovery from meteoroid impact) is a conservative design limit.

40 kW

Bus Motor Control - The Ring bus has 157 identical bays, each with 24 active members. Each member incorporates both thermal and PZ actuators. Budgeting 20 W for each member (not all actuators in all members operate simultaneously) calls for 76 kW. Eight active nodes are allocated to each bay as well. The outboard nodes join 8 members while the inboard nodes join 4 and support the AMCD stations. Since the thermal actuators have been included with the members themselves, budget 20 W for each node, for another 26 kW.

102 kW

AMCD and Mirror Control - Each bay houses a double AMCD suspensor/drive station. Following Appendix A4-5, we assume 125 W per station per rim for both magnetic rim control and PZ mount actuation. The total is 40 kW, to which we add 5 kW for the EM and linear PZ secondary mirror mounts, and tracker drives for sensory intelligence.

45 kW

Nervous System - Including here only sensory devices and the onboard processing network, we follow the biological analogy of Appendix A4-1 to budget 30 % of the power subtotal.

56 kW

As we have done for the other fleet craft, we will superpose these values and inflate the total with a growth margin.

$$\sum_{\text{power}} = \underline{300 \text{ kW}}$$

Mass Estimate

Power Plant - Mass estimates for an out-of-core thermionic reactor producing 300 kWe based on the references in Chapter 9 vary by up to a factor of two. We assume the high end of this range, and increase it further by 50 % because of our extra thick and 4π shielding specification.

15,000 kg

Secondary Mirror Assembly - Here we group the monolithic mirror, its attitude and pointing sensors, its EM space-bearing mounts, the pivot frame and its linear PZ motor, mission (targeting) datonics, and the assembly platform. The assembly is ballasted (for momentum management explained in the next section) to balance the diametrically located power plant.

15,000 kg

Bus Structure - Each bay of the Ring exoskeleton consists of 24 members of various lengths: 4 @ 2 m, 4 @ 4.4 m, 8 @ 3.3 m, and 8 @ 2.6 m. That totals 73 m of member length per bay, or 11.5 km for all 157 bays. Round this up to 12 km since the truss nodes (which require more material) are included, but 10 cm per member (its thermal actuator) is excluded. Assume for reference that the members have the same (hollow circular tube) thickness and diameter as NASA's LEO Space Station [NASA TB, 8705]. Then the volume of material required is $2\pi r t L = 2\pi(.025)(.0015)12(10^3) = 2.83 \text{ m}^3$. The C/Mg composite has a density of 1745 kg/m^3 , so the truss material mass is $(2.83)(1745) = 5000 \text{ kg}$. Assume the aluminum thermal actuator inserts are thicker (for load-carrying) and 10 cm long. Their mass is $2690(157)24(2\pi)(.025)(.002)(.1) = 319 \text{ kg}$.

We will budget 500 kg for the thermal actuators. We also budget 20 % of the member structural mass for its PZ multimorph coatings, and another 20 % extra for the specialized service fittings used to connect the active exoskeleton pieces together. Finally we add 10 % for the fiberoptic strain sensors embedded throughout the structure and applied to its surfaces, and for the fiberoptic temperature sensors included with it.

8000 kg

AMCD - There are two rims, each 11.25 kg/m of their length, like those in the Transducer. Their total mass is $2(11.25)\pi 100 = 7000$ kg. We will specify a greater PZ stroke in the Ring's suspensor/drive stations than the Transducer's can provide. Thus we assume a larger mass (15 kg) per double station. There are 157, for a total of 2355 kg. The AMCDs are enclosed in a square section debris bumper, of 135 kg/m^3 foamed reactive metal, with total material thickness 1 cm. This structure masses $(135)4\pi 100(2)(.01) = 3400$ kg. We allow another 500 kg for dedicated optical sensors in the chase, and 1000 kg for special fittings, and round the total up slightly.

14,400 kg

Propulsion - We avoid the Transducer's extra feedline penalty, but use the same engine mass of 25 kg since the Ring's engines are rated for twice the power. 16 engines are grouped in four clusters of 4. To this 400 kg hardware total (which includes the engines, dry tankage, beam neutralizers, power harnesses and mountings) we add 2600 kg of tanked xenon propellant.

3000 kg

Nervous System - We have accounted for the structure's sensors and actuators elsewhere, so this includes inertial sensors, sensory nerves, processor circuitry, and motor nerves (power distribution throughout the bus). We allocate 20 % of the structure's 8000 kg.

1600 kg

The spacecraft mass subtotal is 57,000 kg, which we should inflate by about 30 % due to the level of uncertainty appropriate for this member of the fleet.

$$\sum_{\text{mass}} = \underline{75,000 \text{ kg}}$$

Inertial Properties

As usual, we choose a cartesian coordinate system with x axis concentric with the bus. We choose the z axis parallel to the line connecting the power plant and secondary mirror assemblies. The y axis makes an orthogonal right-handed triad with the others. Because both the power plant and mirror assembly are on the same "face" of the Ring, the $y - z$ plane does not coincide with the Ring midline. The centroid (coordinate origin) is in fact:

$$\frac{(6.6)39600}{75000} = 3.48 \text{ m away from the Ring plane.}$$

This means of course that rotations about the y and z axes will not merely pivot the Ring about a diametral line. Rather, all rotations will occur about the centroid. Since during operation the Ring remains permanently normal to the Venus - sun line, it practically does not require y or z axis rotations anyway. Such stationkeeping maneuvers as are required must be programmed carefully to permit mirror pointing compensation at the same time. Rotations about the x axis, however, are the basis for the Ring's ability to target the laser anywhere within an entire celestial hemisphere. Since large "rolls" about the x axis are part of normal operation, we balanced the subsystem masses to keep the centroid on this axis, preserving the Ring's inertial symmetry. Thus no extra, propulsive, effort is needed to keep the Ring centered on the Venusian output beam during the rotation maneuvers. To estimate inertial properties, we model the Ring as a thin ring containing the mass of all systems save power, mirror assembly, and propulsion. Those three are considered as point

masses located at their respective centroids. Then:

$$I_{xx} \approx (75,000) 50^2 = 188 \text{ M kg m}^2$$

$$\begin{aligned} I_{yy} &\approx \left[(8000 + 144,000 + 1600) \left(\frac{50^2}{2} + 3.48^2 \right) + \right. \\ &\quad 30,000 (50^2 + 3.12^2) + \\ &\quad \left. 1500 (50^2 + 3.48^2) \right] (1.32) \\ &= 144 \text{ M kg m}^2 \end{aligned}$$

$$\begin{aligned} I_{zz} &\approx \left[(8000 + 144,000 + 1600) \left(\frac{50^2}{2} + 3.48^2 \right) + \right. \\ &\quad \left. 1500 (50^2 + 3.48^2) \right] (1.32) \\ &= 45 \text{ M kg m}^2 \end{aligned}$$

where the second squared term in the parenthetical sets derives from the parallel axis theorem, and the factor 1.32 introduces the 32 % mass inflation factor discussed above, to adjust the tabulated subsystem mass values.

Appendix A4-8 Ring station-keeping performance.

AMCD Authority

The maximum spin rate allowed for the 50 m radius Kevlar AMCD rim is 12.3 rad/s (Appendix A9-4). As usual, we bias the rims about 10 % below their maximum, or at 11.2 rad/s in this case. Each rim therefore stores angular momentum:

$$H = I\omega = mr^2\omega = 3500(50)^2 11.2 = 98 \text{ MN m s}$$

Although 1 cm is a reasonable magnetic gap dimension, we use the PZ suspensor/drive mounts in the Ring to move the location of this gap in a controlled way, thereby gaining an effective gap stroke of ± 3 cm. The maximum angular momentum available for precessional torquing is:

$$H = 2(98)(10^6) \sin \tan^{-1} \left(\frac{.03}{50} \right) = 118 \text{ kN m s}$$

The worst-case angular control rate is about the y axis:

$$\omega = \frac{H}{I_{yy}} = \frac{118(10^3)}{144(10^6)} = 0.82 \text{ mrad/s} = .008 \text{ rpm}$$

While this means that a $\pi/2$ turn would take about 32 min, that rate exceeds the reasonable operational requirement; in order to remain Venus-oriented during the Venusian year, the Ring need only make such a $\pi/2$ turn in 56 d, a rate 2530 times smaller.

Rotation rate about the x axis is set by differential acceleration of the AMCD rims. Even specifying a rate about four times as fast, of 3 mrad/s, which results in a $\pi/2$ turn taking about 9 min, we need only:

$$H = I_{xx}\omega = 188(10^6) (.003) = 564 \text{ kN m s}$$

effected by a differential rim speed of:

$$\omega = \frac{H}{mr^2} = \frac{564(10^3)}{(3500) 50^2} = 64.5 \text{ mrad/s}$$

representing just a 0.29 % speed change for both rims.

Propulsive Capacity

We have budgeted 2600 kg of tanked xenon propellant for the Ring. With no secular drag and no station-exchange maneuvers, the Ring has the lowest propulsive needs of any craft in the fleet, directly related to the accuracy with which L1 and L2 Stations are emplaced at the libration points. Δv of a few m/s per year is appropriate. By choosing an excessive value of 100 m/s per year, we calculate the minimum time between refuelings to be:

$$\begin{aligned} n &= -\ln\left(1 - \frac{m_p}{m_i}\right) \frac{I_{sp} g_0}{\Delta v} = -\ln\left(1 - \frac{2600}{75,000}\right) \frac{3500(9.8)}{100} \\ &= 12 \text{ yr} \end{aligned}$$

for the $I_{sp} = 3500 \text{ s}$ ion engines we have chosen.

Appendix A4-9 Program LINKPERF.FTN.

```
ftn,1,s
$files 2,2
  program linkperf

c communication rate

      real lg,lasr(6),diff(6),rcvr(6),tau(6,6)
      real b(6,6),rs(6),br(6)
      open(99,file = 'ceti.dat')
      open(97,file = 'cetigr.dat')
      pi = 4*atan(1.)

c laser parameters

      etas = .21
      beta = 2.e13
      lg   = 6.e5
      dc   = 1.e3

c diffraction parameters

      snr = 20.
      etad = .84

c receiver parameters

      delr = 5.
c      br
      dr   = 1.e3
c      rs
```

c vary the parameter: target spot size

```
rs(1) = 57.9e9
rs(2) = 108.2e9
rs(3) = 149.6e9
rs(4) = 227.9e9
rs(5) = 778.3e9
rs(6) = 1427.e9
do 2 i = 1,6
```

c vary br

```
do 3 j = 6,1,-1
  br(j) = 1.
  do 4 k = 1,j
    br(j) = br(j) * 10.
4    continue
```

c calculate link data rate

```
lasr(i) = 1. / (etas * beta * lg * dc**2)
diff(i) = ((snr * rs(i))**2) / etad
rcvr(j) = delr * sqrt(br(j)) / (dr**2)
tau(i,j) = ((16./pi) * lasr(i) * diff(i) *
+ rcvr(j))**2
b(i,j) = 1. / (2. * tau(i,j))

write(99,*) tau(i,j),b(i,j),br(j),rs(i)
3    continue
2    continue
```

c generate grafit data columns

```
do 5 m = 1,6
  write(97,10) br(m), (b(n,m),n=1,6)
10  format (7(e8.3,2x))
5    continue

close(99)
close(97)
stop
end
```

References

W A Baracat & C L Butner- [op cit Chapter 9].

Drake Deming & Michael J Mumma- [op cit Chapter 2].

NASA Tech Briefs- (May 1987) [op cit Chapter 9].

F H Schubert, R A Wynveen, P D Quattrone- "Advanced
Regenerative Environmental Control and Life Support
Systems: Air and Water Regeneration" Controlled
Ecological Life Support System NASA CP 2378 (1985).

CHAPTER 5

PLANETARY RESONATORS

Chapter Abstract - A 5-station ring resonator geometry optimizes single-circuit gain, output level smoothness, and satellite number. The sweeping footprint does not contribute significantly to steady-state gain. The Doppler-shift caused by orbiting stations cancels around the ring, permitting oscillation on one CO₂ spectral line. Ring laser gyro theory indicates a bidirectional split-frequency circulating laser field; the system must therefore spoil gain in one direction.

Ring Resonator Configuration

We have already seen (in Chapter 2) that the single-pass gain along tangential paths through the mesospheres of local planets having natural CO₂ lasers may be taken to be about 0.07, if the region of tangency includes the subsolar point. Since to first order the gain in these inverted layers is a linear function of insolation, moving the path's tangent away from the subsolar point in any direction, given planetary spherical geometry, will reduce the single-pass gain to a value depending on the cosine of the solar zenith angle θ_s at that point (Figure 5-1). The cosine function is "tolerant" of moderate excursions from the peak, but a 10% decrement results from a 25° departure, and worsens rapidly beyond that. A resonator configuration useful for operating a planetary laser must therefore maximize approximately subsolar tangency.

The orbital mechanics of a planet-star gravitational system produces only 5 points of zero acceleration relative to the mass line: the Lagrange libration points. We will return to these in a later chapter, but for now we note that the three collinear points could never define a tangent subsolar point, and the two triangular points, located as they are 60° ahead of and behind the planet in its orbit, define a line which does not even intersect the planet. Since an object in any other orbit will move relative to the planet and its atmosphere, our problem reduces to one of finding an orbital satellite configuration which will somehow tap energy from a point which the satellites see only occasionally. Another way of looking at this is to recognize that we must "maximize approximately subsolar tangency" in the time domain.

Suppose for the moment that an orbit can be found which will always pass over the subsolar point, as that point's

orientation changes in inertial space during the course of a planetary year (later we will see two ways of doing this). Then it is instructive to consider the kind of laser output we could get as a function of time. Figure 5-2 shows that output, as available steady-state gain, for a pair of lossless mirror satellites in a circular orbit, spaced such that their line of sight passes through a mesospheric inverted layer (Figure 5-3). Time is measured along the abscissa in terms of true anomaly, such that 0 marks the dawn terminator crossing, and 2π represents a complete orbital period. The gain achieves its nominal 0.07 peak value only fleetingly at noon passage, and of course disappears entirely during the darkside passage.

Now if we add a third satellite, and put them all at the one radius such that their lines of sight define an equilateral triangle tangent at three points in the mesospheric inversion layer (Figure 5-4), we have a ring laser instead of a simple oscillator. That is, the reciprocating oscillatory field is replaced by two counterpropagating circulating fields. Furthermore, one region emerges through the dawn terminator before the previous one plunges through the dusk terminator, so the system gain never drops to zero. Completely analogous to a rectifier circuit, the orbiting ring laser adds the time-varying gains from component segments to fill in each others' gaps, evening out the total gain time profile. In the triangular case, each segment is $2\pi/3$ out of phase with the preceding one, and the sum of all three is quite close to the peak 0.07 (Figure 5-5).

Augmenting the ring with more segments yields even more overlap, so that not only does the variation of the total output (its bumpiness) decrease, the mean value of that total increases beyond the single-pass gain value. Figure 5-6 shows the result for a pentagonal ring; the minimum system steady-state single-circuit gain exceeds by 50 % the planetary

single-pass gain. We may therefore say that the engineered available single-circuit gain is roughly 0.1.

The simple program MARS.FORTRAN which generates these studies comprises Appendix A5-1. Figure 5-7 plots together the system single-circuit gain for lossless resonator polygons up to heptagonal, so that we may compare the configurations. First, we note again that a closed polygon is necessary to achieve a workable duty cycle. Second, although more segments introduce more gain, even-sided polygons produce bumpier envelopes than do odd-sided polygons. This makes sense if we consider that diametrically opposed satellites are by definition π out of phase, so that one segment's contribution dies away to zero before its diametric replacement comes on line. Odd configurations have no opposed satellites, so segments' arrivals into and departures from sunlight never coincide.

Smoothness of available gain is desirable for nominal operation of a communication laser. In Chapter 3 we saw that one source of degraded system signal-to-noise ratio is a non-zero extinction ratio; having a non-steady laser output to begin with simply multiplies the probability of recording a bit erroneously upon receiving the signal. Furthermore, while a receiver sophisticated enough to record modulated interstellar signals properly could also track regular variations in maximum signal strength, minimizing those excursions as much as possible at the transmitter would naturally simplify that task, improving the quality of effective data transfer. Figure 5-7 shows immediately a preference for odd-sided polygons. The available gain envelope for the pentagon, for instance, varies only about 7 % peak-to-peak.

There is some reason to believe that the variations will not be as mathematically cusped as shown in Figure 5-7. First, the mesospheric gain does not decrease as much as a strict cosine dependence would predict [Deming & Mumma, 83]. Because of the gain region's altitude, the population inversion continues slightly beyond the planetary terminator. This additional overlap will tend to smooth further the first-order envelope shown in Figure 5-7. Additionally, discussions with Mumma [87] indicate that real lasers tend to equilibrate their own operation such that abrupt changes in such basic features as circulating field strength become evened out. Thus in the case of an odd-sided planetary ring laser, we would expect a smoothly rippled gain envelope. Our reference Transducer can neutralize the far-field effect of this small ripple by varying its defocus bias in a programmed way.

Since the available gain is seen to increase so dramatically with an increasing number of satellites, why not just keep adding more? In the limit, we could propose a continuous reflective band encircling the planet, comprising what is known as a "whispering gallery" laser cavity. This would extract the maximum possible energy from the mesospheric inversion. Of course, it would also shade the gain region from the sun which pumps the laser! For a variety of reasons (ranging from oscillator mode selection to the dynamic stability of a thin wide structure suspended in an atmosphere with tethers from a higher orbit!) we will regard that limit as of little practical interest for this study.

Returning to the discrete satellite arrangement, there are three basic reasons why wantonly adding more becomes a liability. First, in a system as complex as this one already has to be, complexity (and attendant expense) which is not strictly necessary is unwarranted. Second, as the number of sides increases, the polygon more closely approximates the

circle defined by the mesospheric inversion layer. Thus more satellites mean a lower orbit, eventually making both the continuous propulsive effort required to overcome atmospheric drag, and the sustained systems degradation incurred by atomic fluxes in the planetary exosphere, not worth the increased laser gain. Third, no planet is a perfect sphere, and the higher-order deviations from sphericity produce a gravity field which varies greatly and abruptly at low altitudes, yanking an orbiting satellite away from its nominally Keplerian orbit. Since the mesospheric inversion layer is typically only about 10 km thick, not much yanking can occur before the satellite line of sight misses the laser medium altogether. Also, the lower the satellites orbit, the faster they must move, so worsening gravity anomalies must be compensated more quickly, on a scale of minutes or even seconds. These considerations indicate that fewer satellites, farther away from the planet, are better.

On the other hand, we cannot design a communication system based on extractable energy greater than the minimum of the gain envelope, which in the case of the triangular geometry is about 85 % of the single-pass peak, for a single-circuit gain of less than 0.06. Considering our currently imprecise knowledge of the single-pass gain value of 0.07 (Gordiyets and Panchenko's simpler model [82] says ~2 - 40 % for Venus!), and knowing that the system will have losses which must be subtracted from that value, choosing three satellites leaves little margin. It would seem prudent to take some advantage of the engineered available gain improvement just discussed. Furthermore, the triangular geometry precludes reflection incidence angles greater than 60°, which does not bode well for minimizing reflection losses.

This project presumes 5 satellites, in the pentagonal resonator geometry, as its baseline configuration for planetary

lasers. We will see how the reflection incidence angles in this case can be made 72° , approximating normal incidence conditions (to within a 5 % cosine tolerance). If more detailed planetary measurements should subsequently reveal a single-pass gain enough smaller than 0.07 to require more resonator segments, overall system changes from our pentagonal baseline would be more minor than from a triangular design.

Sweeping Footprint

The single-circuit gain accomplished by a ring resonator in the previous section is called "steady state". The calculation presumes that the active region --- that point on the resonator line of sight tangent in the inverted population of CO_2 molecules --- does not move relative to the atmosphere. Rather, the laser is considered to extract energy from the same volume of gas for times long compared to the transition lifetimes of the molecule, so that energy can be taken out only as fast as the sun pumps it in.

Some prior, concept studies of planetary lasers [Britt, 82] have suggested an improvement over the steady-state gain, resulting from a "sweeping footprint". This refers to the fact that an active region defined by orbiting stations will move around the planet with the same angular velocity as the satellites, sweeping through a volume of inverted atmosphere greater than the static assumption would calculate. The undepleted mesosphere encountered would add energy to the laser at greater than steady-state rates. Since the argument typically yields order-of-magnitude gain improvements, we will examine it in detail for the two-station case. Figure 5-8 shows this argument graphically.

Its fallacy is the orthogonality of the orbital path and the laser beam. Given a physical separation of the mirror stations, particularly one large enough to keep them out of the exosphere, this geometry would require them to be moving in exactly parallel orbits (Figure 5-9) --- that is, non-planet-centered orbits! Now, we can invent fantastic ways of doing this. They could be held apart by a structure thousands of kilometers long, which of course would drag through the atmosphere itself. Or solar sailing techniques could hold them away from their nominal orbit, except on the darkside, since planetary eclipse consumes a substantial time fraction of low orbits. Or, if the satellites had unbelievably low mass, the circulating power of the laser field itself would keep them apart, except of course during the darkside passage, when they would accelerate back toward a rendezvous at their proper orbit. Appendix A5-2 presents a consumable propulsive method of achieving this configuration, which demonstrates the utter impracticality of non-Keplerian orbits.

All such schemes leave untouched the more advanced problem that their hard-won gain regions would still result in a gain envelope of the type shown in Figure 5-2, despite a higher peak value. Thus whatever solution were found to the perturbed-orbit problem would need to be duplicated several times (5, for instance) and co-orbited. The separate outputs would then be combined in phase at some (moving?) transmitter station to yield a useful communication duty cycle. The simplest approach, stationing the satellites in intersecting planet-centered orbits and allowing them to approach and recede from each other (Figure 5-10), essentially squares the duty cycle dilemma --- not only do the satellites move around the planet away from the subsolar point, their line of sight changes in altitude as their separation changes. This configuration would insure getting only a brief peep of laser

light out of the system on each orbit (although that peep might be a real blast). The usefulness of such a beacon for communicating would be analogous to that of the rotating light on a police car.

This study will consider further only the more reasonable geometry in which the laser beam propagates parallel to the resonator orbit plane. The only effect of the sweeping footprint in this case is to elongate the active region. To see what benefit this extra length yields, we will use as a numerical example this chapter's actual design. At Venus, the stations comprising a pentagonal resonator orbit at a circular speed of 6.5 km/s. The active region accordingly advances with circular speed 5.3 km/s (Venus' equatorial atmosphere can be considered to move with a mean velocity equal to the circular speed of the planet itself, which at roughly 2 m/s is insignificant). It takes light about 0.15 s to complete one circuit of the resonator, in which time the active region has advanced about 800 m. Thus the "fresh" mesosphere encountered per circuit corresponds only to about 0.001 of the active region's nominal length. The gain available from this undepleted gas would have to be immense indeed for it to make any difference. We therefore can ignore any additional gain introduced by the beam's sweeping footprint.

Doppler Shift

Electromagnetic energy which originates from or is received by a moving object is subject to the Doppler shift, which changes its frequency. This may be understood by noting that in the time period required for one wavelength, the location of a moving source or receiver will have changed

(Figure 5-11), so that the wave's nodes will be forced closer together (upon approach) or farther apart (upon retreat) spatially; its time frequency thus respectively increases (blueshifting) or decreases (redshifting). For a combined source - receiver system, the shift can be calculated as arising from a change in path length over time, dl/dt .

An orbiting planetary laser relay station can for the moment be treated as a "black box" which receives the beam from the previous station via the mesosphere, and redirects it to the subsequent station, again via the mesosphere. To clarify the illustration, we show it as a simple plane mirror in Figure 5-12. The station's orbital circular velocity at any point can be broken down as shown into components parallel and perpendicular to either the incoming beam or the outgoing beam. Only the parallel components affect the light's frequency, and from symmetry the incoming and outgoing parallel components must have the same magnitude. Consider a laser beam originating in active region A with frequency ν_0 . The orbital relay station, receding from A, will see the beam as redshifted to $\nu_0 - \Delta\nu$. That station approaches active region B, however, so that B will see the station's light (now at frequency $\nu_0 - \Delta\nu$) as blueshifted by the same amount $\Delta\nu$. Thus $\nu_0 - \Delta\nu + \Delta\nu = \nu_0$, and B sees the light with its original laser frequency ν_0 . Looking at the entire ring, we see that although the optical path length is longer than if the satellites were not moving (since the "first" mirror has moved by the time the light gets around to it again), that longer length itself does not change with time for a constant orbital velocity; $dl/dt = 0$. The Doppler shift caused by orbiting the mirrors cancels in each segment and does not accumulate with successive circuits.

That cancellation is critical. We have seen in Chapter 2 that the CO₂ emission spectrum comprises a "picket" structure

of narrow (~50 MHz for the Doppler-broadened emission of the Venus mesosphere), evenly-spaced (~40 GHz apart) peaks due to the quantized ro-vibrational energy levels allowed the molecule. The first-order Doppler shift from pentagonal orbital velocity at Venus (Appendix A5-3) seen by any given mirror satellite amounts to 500 Mhz. If this were not cancelled, the laser would "walk off" the proper gain profile with just one reflection, and oscillation would never have a chance to build up a stable resonant field in the ring cavity.

Ring Laser Theory

Our earlier observation that the actual circuit length around the resonator depends on the orbital speed of the ring indicates that by engineering a vast rotating ring laser, we have touched the topic of ring laser gyroscope (RLG) rotation sensors. We must therefore look at the problems typically inherent in such devices, to evaluate their relevance for this project.

Ring laser gyroscopes (Figure 5-13), which now are increasingly common in at least commercial aviation [Sargent, 84], have been difficult to perfect but worth the trouble because of some advantages over precision mechanical gyros. Specifically, they are lighter, cheaper, have no start-up time, and are reliable for several tens of thousands of hours of continuous operation [Chow et al, 80]; advanced models have no moving parts. They depend on the Sagnac effect, first studied in 1913, which uses rotation to split the frequencies of counterpropagating waves in a resonator [Sargent, 84].

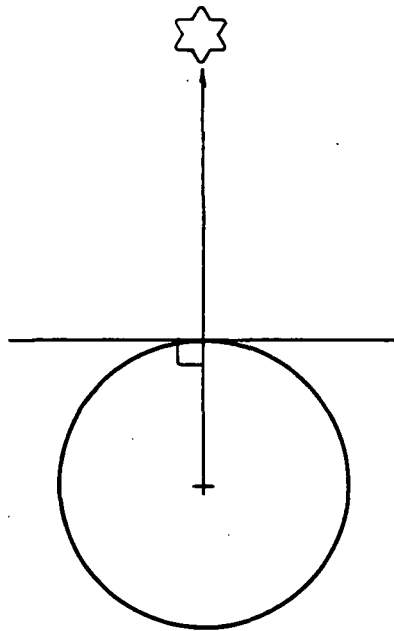
Classically, the rotating ring (Figure 5-14) presents different path lengths, and hence different resonant frequencies, to the waves travelling in opposite directions since the resonator will have moved through a non-zero angle ω in the time τ it takes light to go around [Ezekiel & Arditty, 82]. When coupled out and interfered, the two waves generate a moving fringe pattern, measurable with photodetectors, the velocity of which is proportional to the angular velocity normal to the resonator plane. This Sagnac interference detection is a much more sensitive method for measuring acceleration than the Doppler shift [Post, 67]. The ring laser equations can only be rigorously derived using general relativity, since the light propagates in a non-inertial frame, but higher-order subtleties due to frame dragging and the relative motion of an intra-resonator medium are academic only [Post, 67]. Jacobs & Zanon [82] provide a proof based on Stokes' theorem which generalizes the equations to resonators of both arbitrarily closed shape and rotation axis locus.

The problems plaguing ring laser gyros can be divided into three categories, illustrated in Figure 5-15 as deviations from the linear rotation rate - fringe velocity dependence. Null-shift error is a non-zero measurement despite zero rotation, while scale factor error is a departure from the simple mathematical expression relating frequency difference to rotation rate; neither their causes nor their presence need concern us, since they are important only in cases where variable rotation is measured. However, lock-in is a physical property of rotating resonators and must be addressed.

Homogeneously broadened (mainly solid state or compressed gas) laser media are useless for ring laser gyros because they support only bi-stable, unidirectional fields. Instead RLGs always use He-Ne lasers, with a 1:1 mixture of Neon isotopes whose line centers are naturally separated by about 800 Mhz

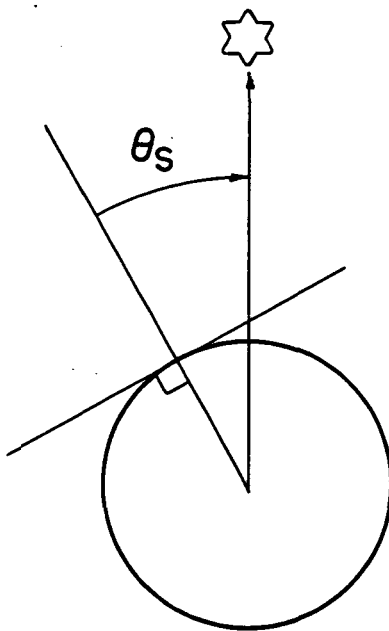
[Sargent, 84]. Tuning the resonator between these frequencies allows the counterpropagating waves to draw from different atomic populations in the active medium, thus avoiding saturation-coupling. The waves can still become phase-coupled if their frequencies differ by only a small amount, and lock to the same frequency, a phenomenon common in many physical systems. The mechanism for this is that energy from one resonant mode is scattered by the imperfect resonator mirrors into the other mode. Since lock-in results in a "dead zone" of slow rotations which cannot be measured, ring laser gyro designers have adopted extraordinary measures to reduce the problem, including high-rotation-rate biasing, high-frequency mechanical dithering (vibrating), polarized wave separation (DILAG), magnetically-induced Faraday detuning (ZLAG), and multioscillator lasers, in which completely separate lasers are made to share the same cavity [Chow, 80].

Ring laser gyro theory is relevant for rotating planetary lasers because the low-pressure CO₂ laser, not exempted from the Sagnac effect as are the homogeneously broadened media, can support counterpropagating, split-frequency waves. We would prefer to have oscillation on only one line of the CO₂ spectrum (to ensure the oscillation threshold), and to have that line be as narrow as possible (to maximize the link signal-to-noise ratio). If we can apply ring laser gyro theory to the enormous scale of planetary lasers, Appendix A5-4 shows that we do not expect lock-in, but that the frequency difference between the counterpropagating modes is about 1/40 the CO₂ line spacing, and about 200 times the linewidth. To first order, therefore, we would expect two distinct counterpropagating frequencies, drawing energy through the same CO₂ spectral line. The detailed optical path design of Chapter 7 provides a way of avoiding this frequency split while simultaneously coupling a useful fraction of the intracavity field out of the planetary resonator.



$$\theta_S = 0$$

$$\cos \theta_S = 1$$



$$\theta_S \neq 0$$

$$\cos \theta_S < 1$$

Figure 5-1 Cosine dependence of single-pass gain.

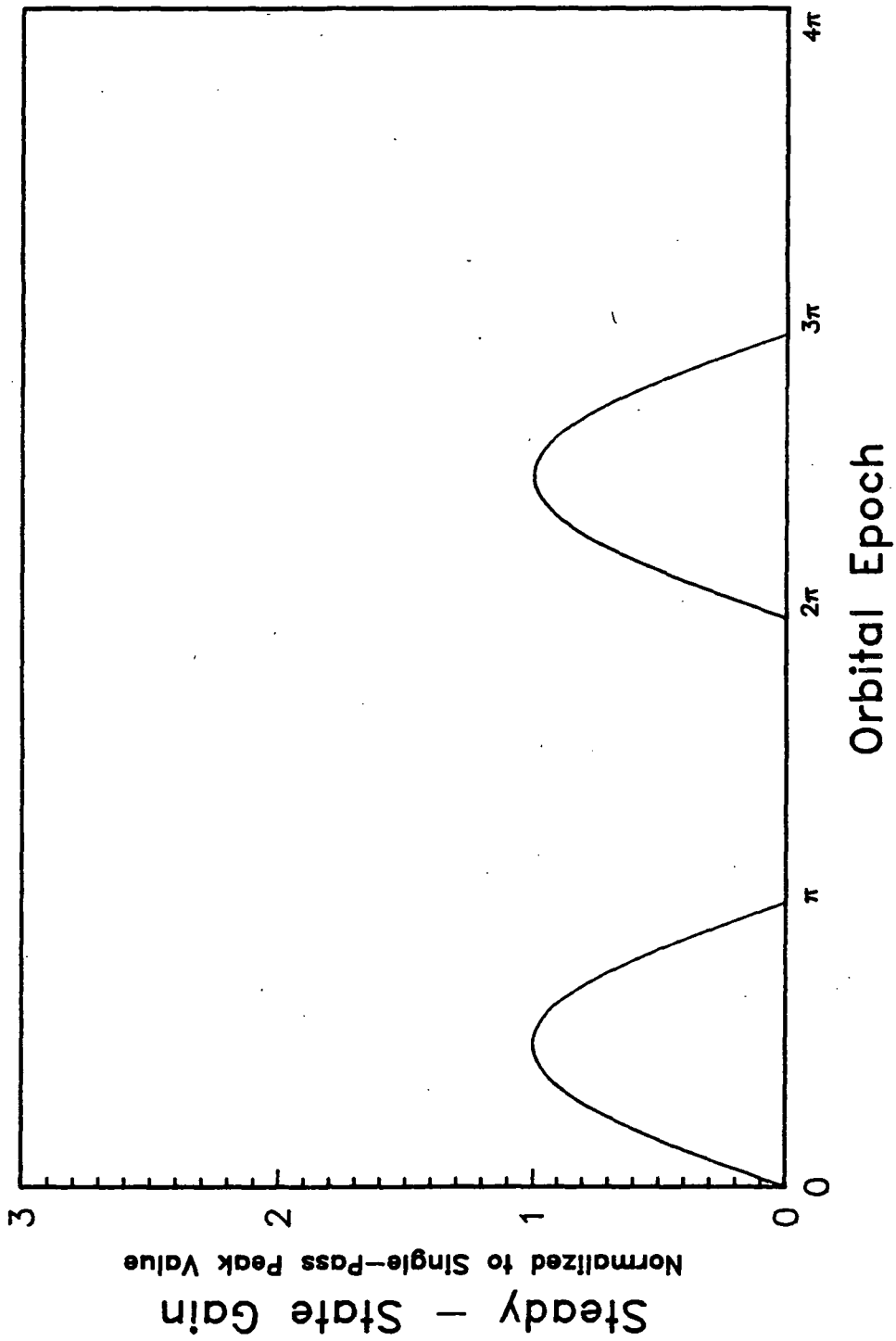


Figure 5-2 Available steady-state gain for a pair of lossless orbiting resonator mirrors.

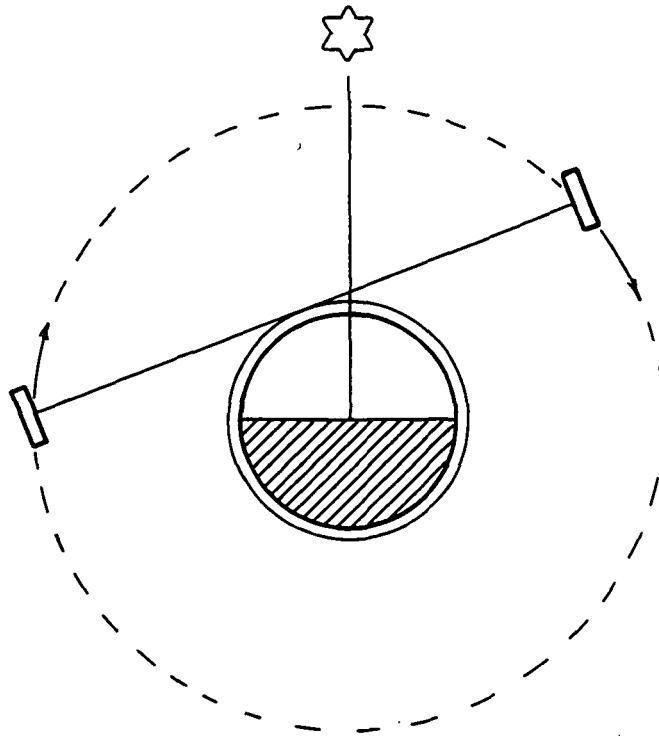


Figure 5-3 An orbiting pair of satellites defining a simple resonator.

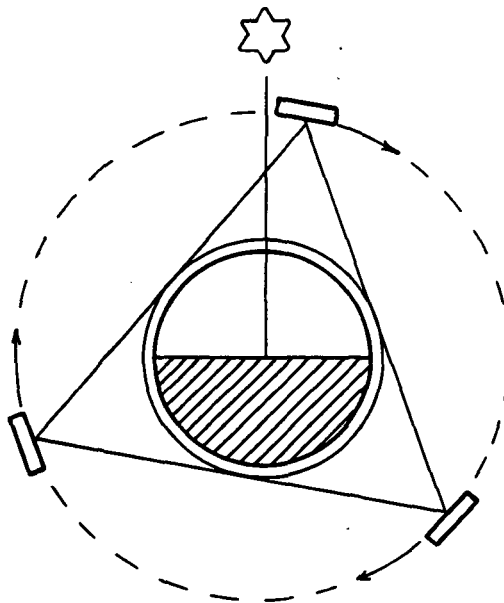


Figure 5-4 Three satellites defining a triangular ring resonator.

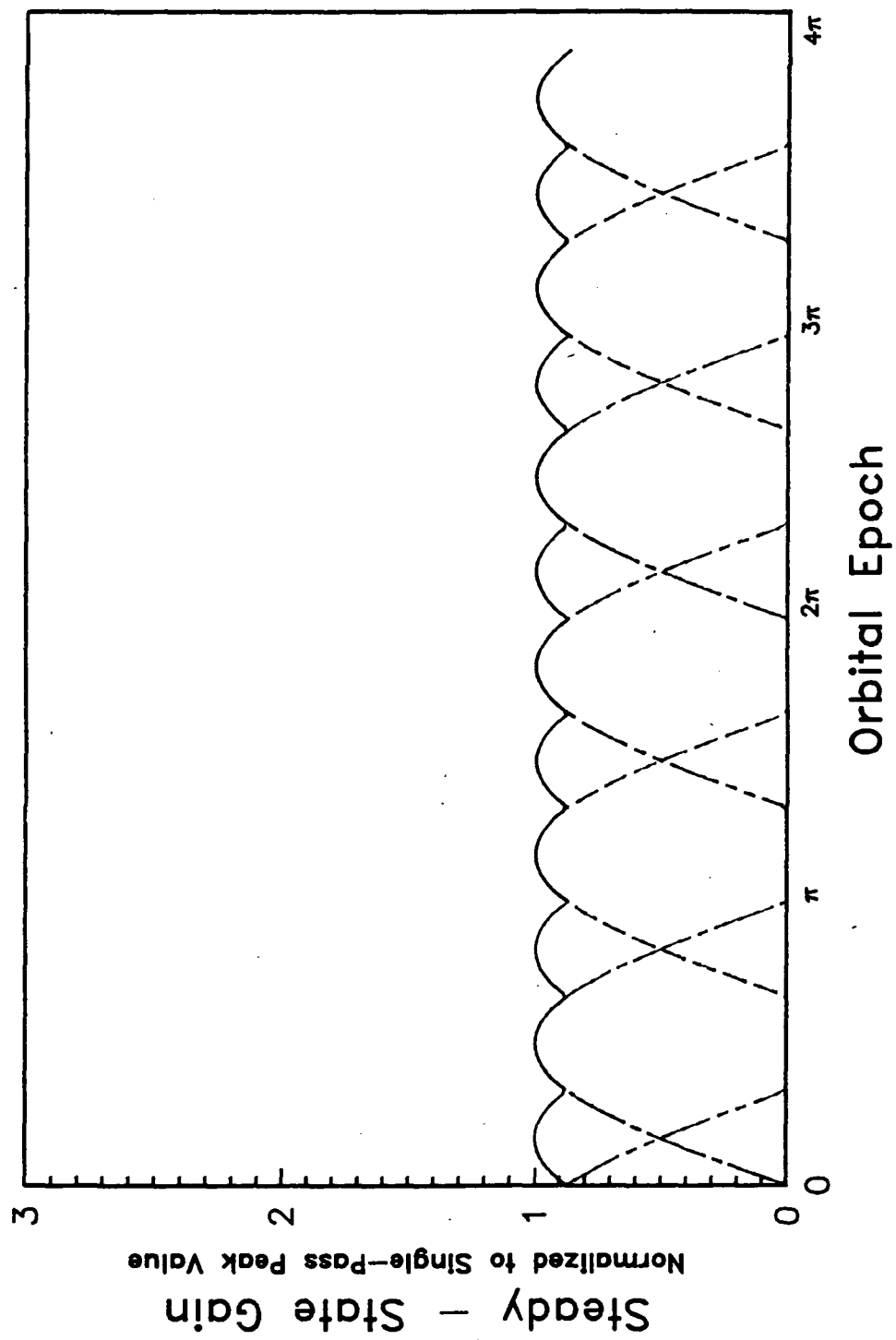


Figure 5-5 Available steady-state gain for a lossless triangular planetary resonator.

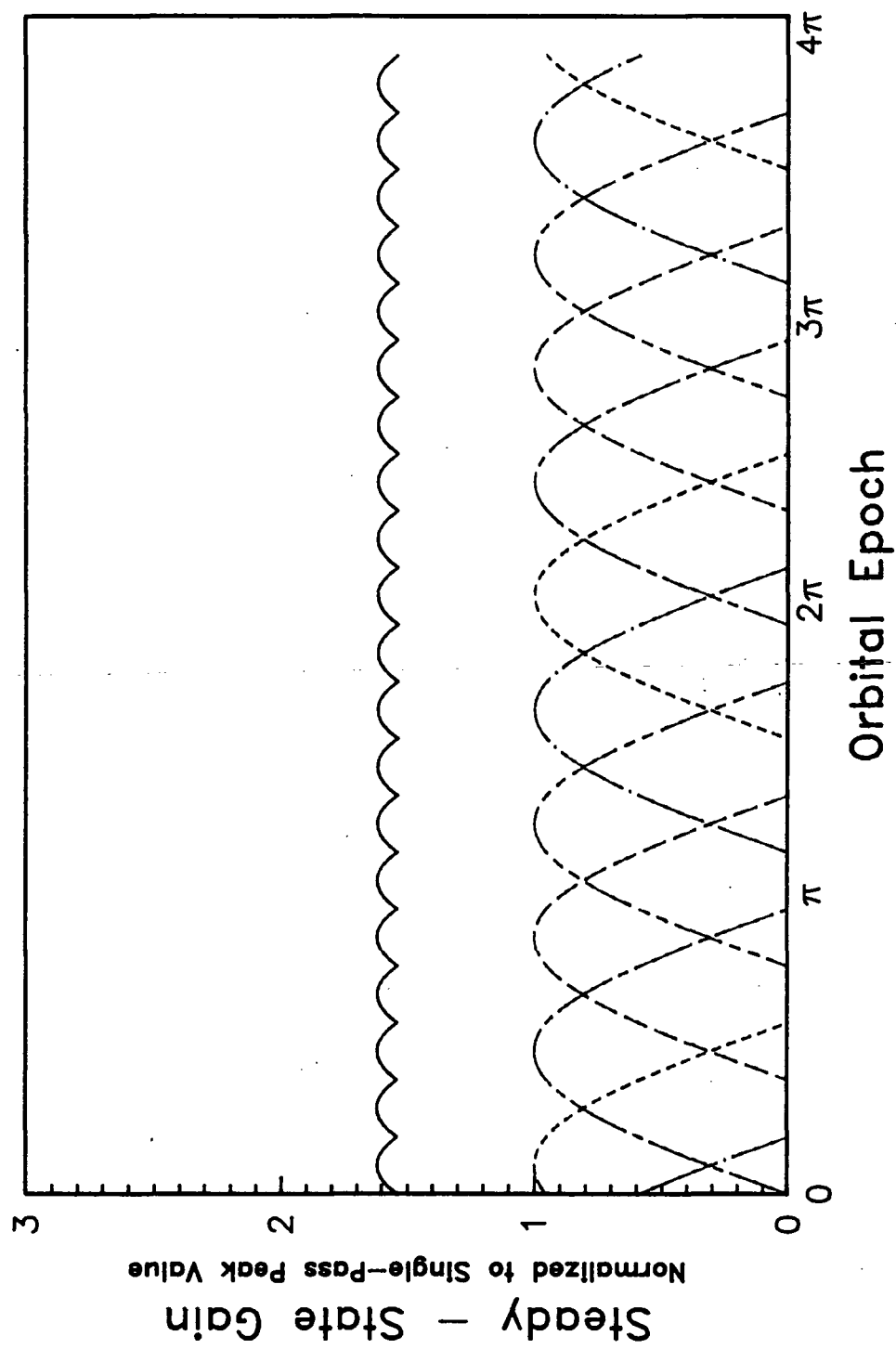


Figure 5-6 Available steady-state gain for a lossless pentagonal planetary resonator.

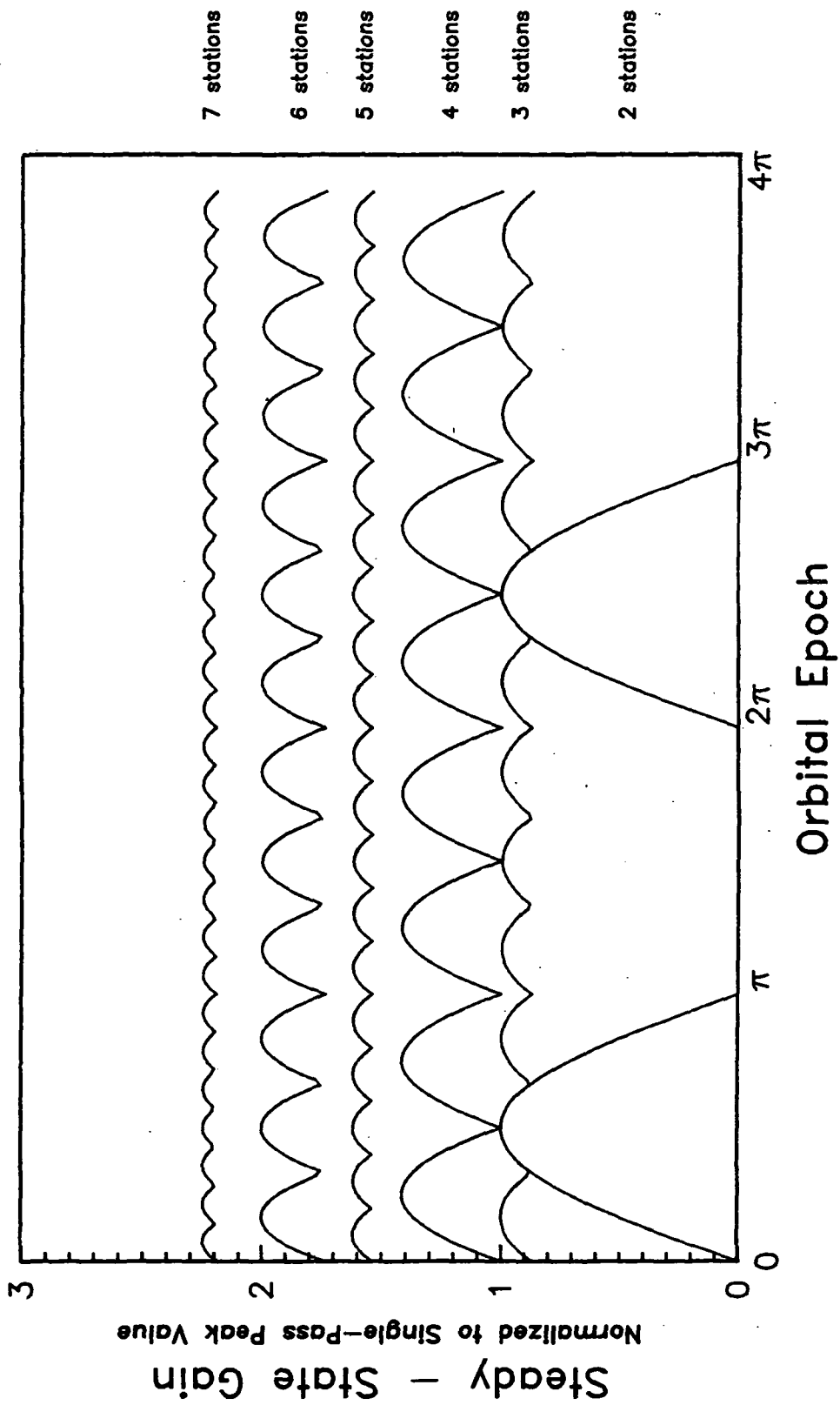


Figure 5-7 Comparison of available single-circuit steady-state gain for lossless planetary resonator polygons up to heptagonal.

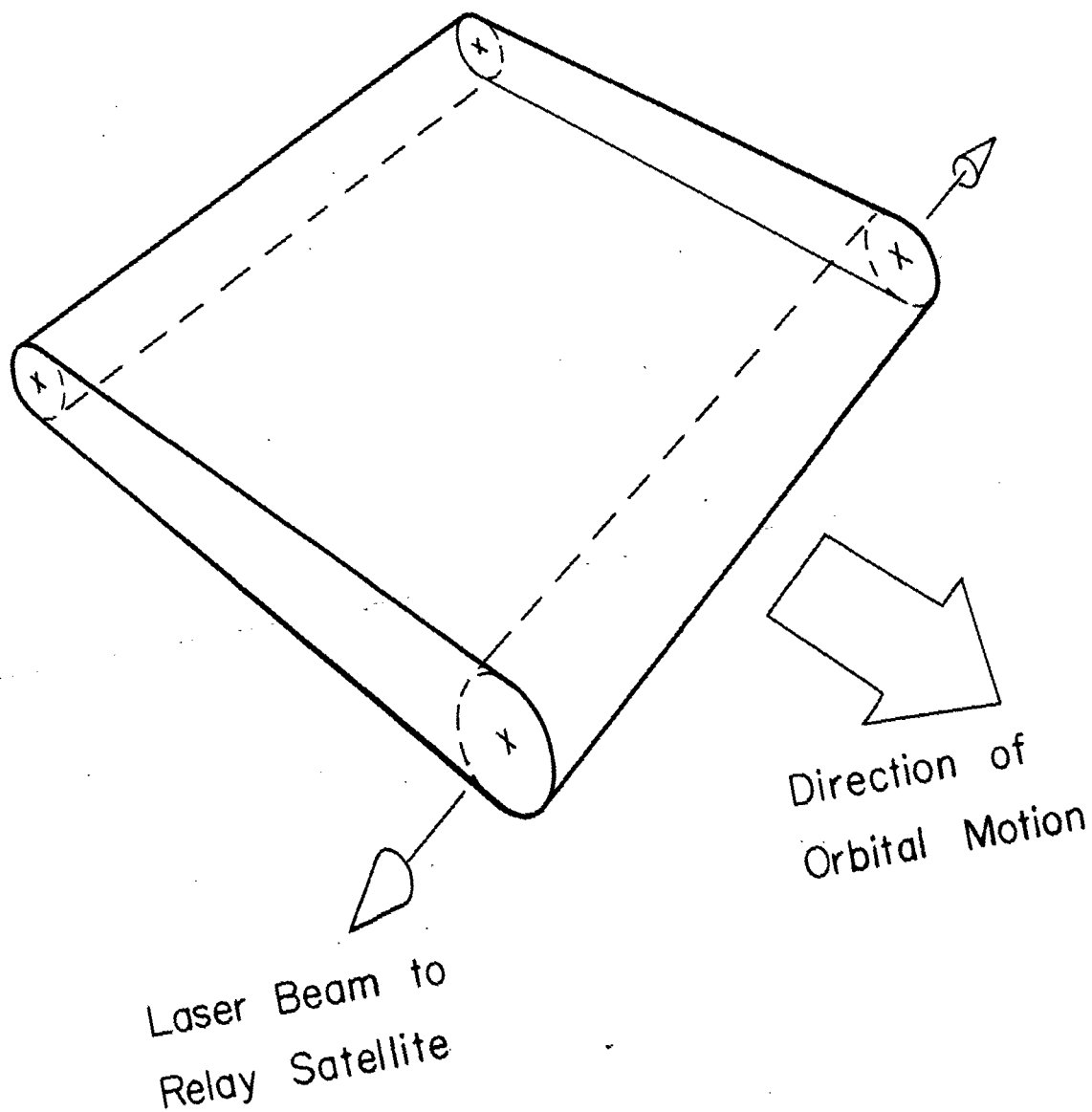


Figure 5-8 Geometry of the sweeping footprint.

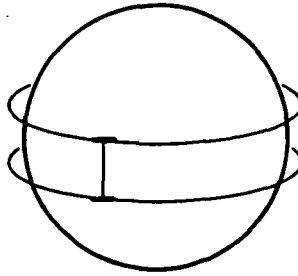


Figure 5-9 Parallel non-planet-centered orbits capable of yielding the orthogonally sweeping footprint.

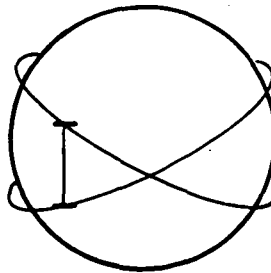


Figure 5-10 Crossing planet-centered orbits yielding a transient gain path.

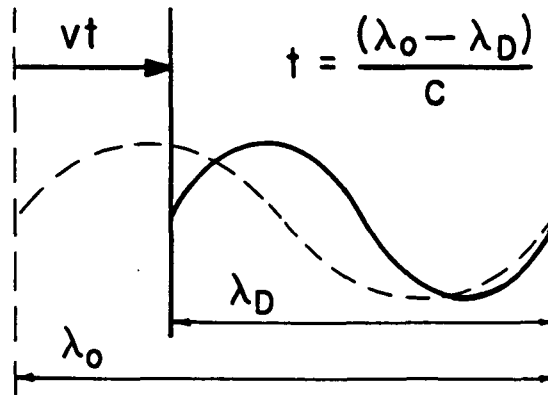


Figure 5-11 A physical explanation for the origin of Doppler shift.

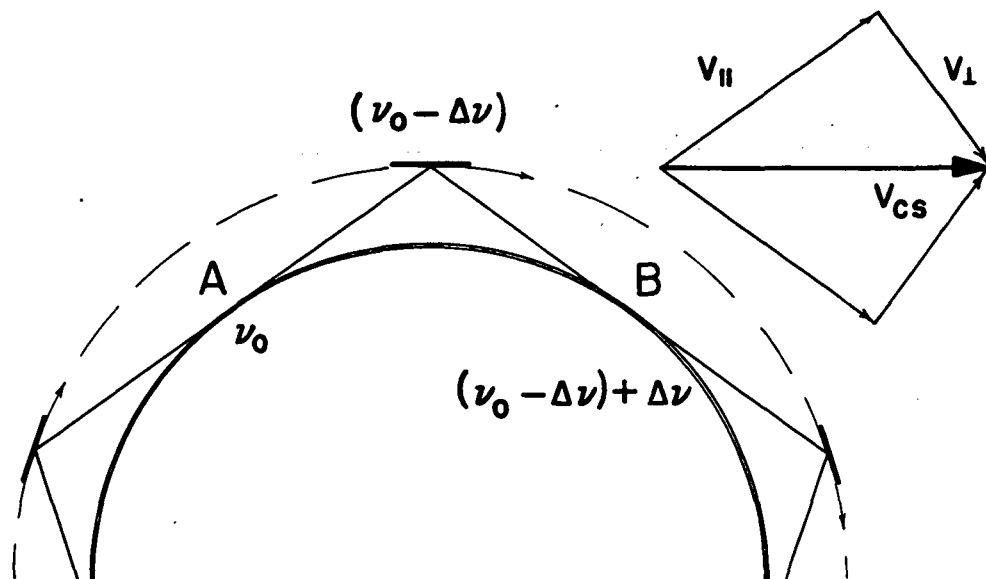
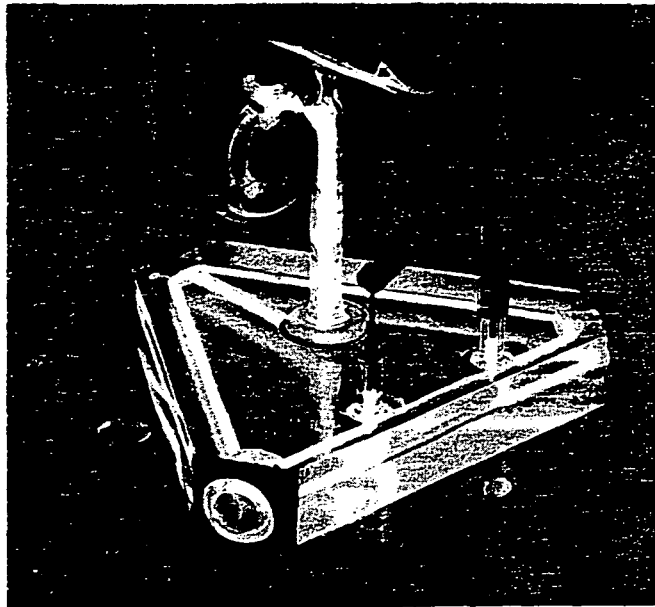


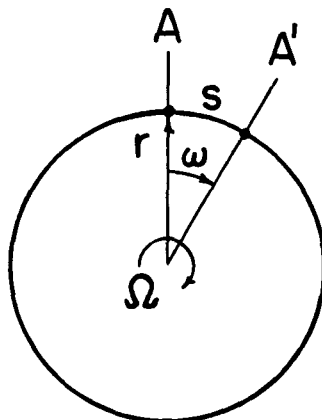
Figure 5-12 The net cancellation of Doppler shift in a planetary laser beam redirected by an orbiting satellite.

ORIGINAL PAGE IS
OF POOR QUALITY



A ring laser gyro. [Photo courtesy Autonetics, a Division of North American Rockwell Corp.]

Figure 5-13 A commercially available ring laser gyroscope. [Hecht & Zajac, 74]



$$\begin{aligned}
 \Delta L &= 2s \\
 &= 2r\omega \\
 &= 2r\Omega\tau \\
 &= 2r\Omega\frac{2\pi r}{c} \\
 &= 4(\pi r^2)\frac{\Omega}{c} \\
 &= \frac{4A\Omega}{c}
 \end{aligned}$$

Figure 5-14 The physics of a rotating ring resonator, showing that the difference in path lengths ΔL for counterpropagating fields varies with enclosed area A and rotation rate Ω .

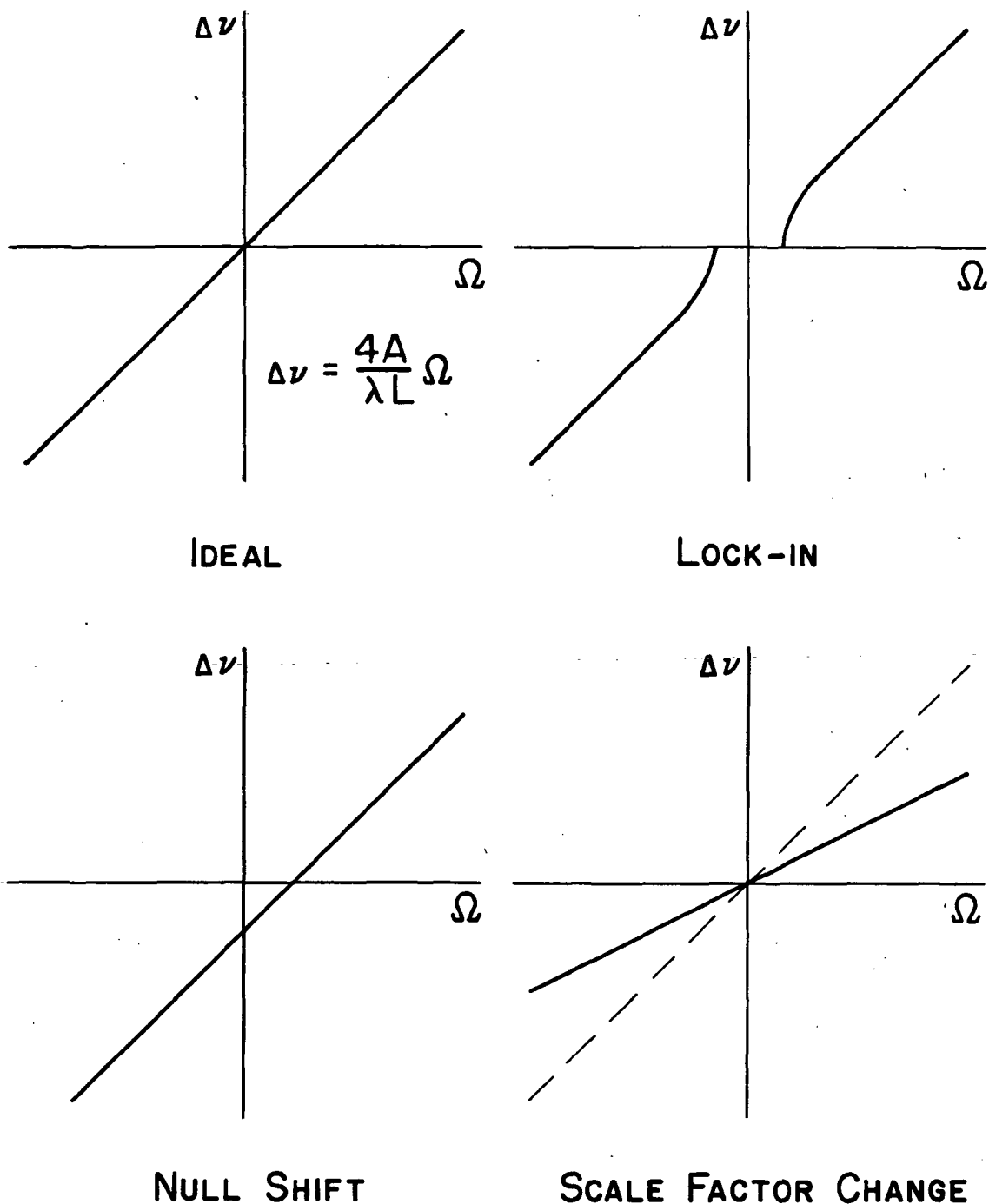


Figure 5-15 Types of ring laser gyro error plotted as deviations from the ideal dependence of beat note (frequency difference) on gyroscope rotation rate. After [Chow et al, 80]

Appendix A5-1 Program MARS.FTN for calculating lossless available single-circuit gain of planetary resonators.

This program simply generates a number of sine curves equal to the specified number of resonator sectors (7 in this example), spacing them evenly within a 2π interval. After eliminating the negative parts of these functions, it adds them to yield available gain. The numerical step is $\pi/250$, and 1000 such steps are calculated, showing the gain envelope for 2 full orbits.

```
ftn,1,s
$files 0,1
      program mars
      dimension x(8)
      open(99,file = 'mars.dat')
      pi = 4*atan(1.)
      ti = pi/250.
      n = 7
      t = 0.
      do 3 k = 1,1001
        y = 0.
        do 2 j = 1,n
          i = j - 1
          x(j) = sin(t - 2*pi*i/n)
          if (x(j).lt.0.5) then
            x(j) = 0.
          endif
          y = y + x(j)
2        continue
        write(99,'(9(f5.3,2x))') t, (x(j), j = 1,n), y
        t = t + ti
3      continue

      close(99)
      stop
      end
```

Appendix A5-2 An argument against propulsively maintaining
orthogonally sweeping gain.

Estimate satellite mass:

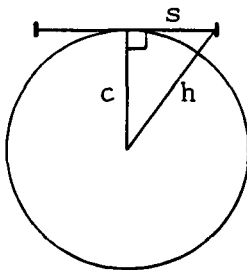
Assume the overall mirror diameter ≈ 2 km, comprised of individual beryllium segments (Chapter 7), each of diameter 1 m and mass 30 kg.

$$\text{Then } \frac{\pi R^2}{\pi r^2} = \frac{(1000)^2}{(0.5)^2} = 4(10^6) \text{ segments will make up the mirror,}$$

which will have a total mass of $4(10^6)30 \approx 10^8 \text{ kg} = 10^5 \text{ MT}$.

Because of substantial distributed structure, sensor and actuator hardware, and other spacecraft housekeeping needs, assume the total spacecraft mass is roughly twice the mirror mass, or $2(10^5) \text{ MT}$.

Choose satellite separation:

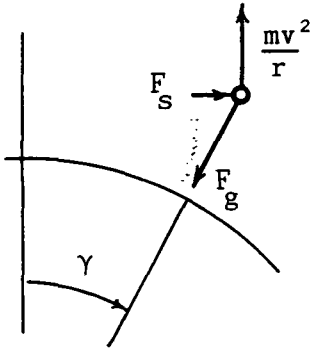


The satellites can be thought of as being perturbed laterally from a common nominal orbit which passes through the tangent gain region. For Venus, c is the sum of the planetary radius (6052 km) and the gain layer altitude (130 km). h must be ≥ 7000 km, because an orbital altitude < 1000 km would

introduce challenging stationkeeping perturbations due to planetary gravity anomalies (Chapter 6).

$$\text{Thus the half-separation } s = (7000^2 - (6052 + 130)^2)^{\frac{1}{2}} \approx 3300 \text{ km.}$$

Separation force required:



The free-body diagram shows the forces acting on the perturbed satellite as it crosses through the plane of the page. The gravitational force F_g is calculated first:

$$F_g = \frac{\mu_g m}{h^2} = \frac{3.257(10^5)(10^8)(10^3)}{(7000)^2} \left(\frac{\text{kg m}}{\text{s}^2} \right) = 6.6(10^8) \text{ N}$$

The separation force which must be supplied propulsively is simply the lateral component of F_g :

$$F_s = F_g \sin \gamma = F_g \frac{3300}{7000} = 6.6(10^8) \left(\frac{33}{70} \right) \approx 3(10^8) \text{ N} = 300 \text{ MN.}$$

Propulsive requirement:

Since the thrust needed to keep the satellites apart is constant over extremely long mission times, assume ion engines. Although they have low thrust, they can be ganged together, operated continuously for many thousands of hours, and depended on for the highest efficiency. From the definition of specific impulse:

$$\dot{m} = \frac{T}{I_{sp} g_o} = \frac{3(10^8)}{3000(9.8)} \left(\frac{\text{kgms}^2}{\text{s}^2 \text{sm}} \right) \approx 10^4 \frac{\text{kg}}{\text{s}}$$

This means that even the most efficient rocket engines known would have to expel an amount of fuel equal in mass to the satellite itself every six hours! Clearly this is impractical; furthermore, all of this propellant would be expelled in the direction of the facing satellite and its optical surfaces. This short calculation and its extreme result show quantitatively just how inarguable the reality of astrodynamics is.

Appendix A5-3 First-order Doppler shift experienced by a laser beam relative to one relay satellite of a planetary resonator.

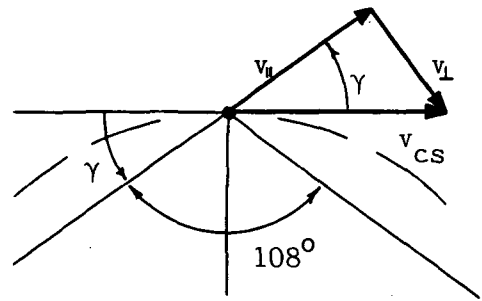
The apex angle of a pentagon is 108° .

See again Figure 5-12. The component

$$\begin{aligned} v_{\parallel} &= v_{CS} \cos \gamma \\ &= 6.54 \cos(36^\circ) = 5.29 \frac{\text{km}}{\text{s}} \end{aligned}$$

Light with wavelength $10.6 \mu\text{m}$ has

$$\begin{aligned} \text{frequency } \nu &= \frac{c}{\lambda} = \frac{3(10^8)}{10.6(10^{-6})} \\ &= 2.83(10^{13}) \text{ Hz.} \end{aligned}$$



Upon redirection from the moving satellite, then, the light's frequency

$$\text{is shifted by } \Delta\nu = \nu \frac{v_{\parallel}}{c} = 2.83(10^{13}) \frac{5290}{3(10^8)} = 5(10^8) \text{ Hz} = 500 \text{ Mhz}$$

Appendix A5-4 The ring laser lock-in threshold applied to the pentagonal planetary resonator at Venus.

The lock-in threshold for a rotating ring resonator [Aronowitz, 71] is calculated by quantifying the energy scattered by a resonator mirror. Assume mirrors of reflectivity $r = .995$ (Chapter 7), so the scattered fraction $r_s = .005$. Assume a laser beam diameter of 2 km. Calculate the enclosed ring area A by partitioning the pentagon into 10 equal right triangles. Each has altitude equal to the sum of Venus' planetary radius and the gain region altitude, or 6182 km, and base equal to the product of that altitude and the tangent of the central angle, 36° . Thus:

$$A = 10\left(\frac{1}{2}(6182^2)\tan 36^\circ\right) = 1.4(10^8) \text{ km}^2$$

The lock-in threshold Ω_L is given by:

$$\begin{aligned}\Omega_L &= \frac{c\lambda^2 r_s}{32\pi A d} = \frac{3(10^8)(10.6(10^{-6}))(.005)}{32\pi 1.4(10^{14}) 2000} \\ &= 6(10^{-24}) \frac{\text{rad}}{\text{s}}\end{aligned}$$

The rotation rate Ω of the orbiting ring itself is (Appendix A6-7):

$$\Omega = \frac{2\pi}{7354} = 8.54(10^{-4}) \frac{\text{rad}}{\text{s}}$$

Since $\Omega \gg \Omega_L$ the resonator will not frequency-lock, and can support two independent cavity modes. The perimeter p of the ring is simply 10 times the base of a component triangle described above, and the frequency separation of the cavity modes is:

$$\begin{aligned}\Delta\nu &= \frac{4A\Omega}{L\lambda} = \frac{4 \cdot 1.4(10^{14})}{10(6182\tan 36^\circ)10^3 \cdot 10.6(10^{-6}) \cdot 7354} \cdot 2\pi \\ &\approx 1 \text{ GHz}\end{aligned}$$

This frequency difference is 200 times larger than the 50 MHz emission width, and 1/40 of the 40 GHz line spacing.

References

- Frederick Aronowitz- The Laser Gyro Vol 1
(Academic Press, 1971).
- Edward J Britt- "Commentary on the JASON Report" Rasor
Associates, personal communication (11 Mar 1982).
- J Chamberlain, W Happer, J Katz, R Novick- Long Path Solar
Pumped Lasers Jason Committee Technical Report JSR-81-07
(Sep 1981) [additional commentary memorandum by Deming &
Mumma].
- W W Chow, J B Hambenne, T J Hutchings, V E Sanders, M Sargent
III, M O Scully- "Multioscillator Laser Gyros" IEEE
Journal of Quantum Electronics Vol QE-16 No 9
(Sep 1980).
- Drake Deming & Michael J Mumma- [op cit Chapter 2].
- S Ezekiel & H J Arditty- "Fiber-Optic Rotation Sensors:
Tutorial Review" Fiber-Optic Rotation Sensors and Related
Technologies (Springer-Verlag 1982).
- B F Gordiyets & V Ya Panchenko- [op cit Chapter 2].
- Eugene Hecht & Alfred Zajac- Optics (Addison-Wesley, 1974).
- Stephen F Jacobs & Ray Zanoni- "Laser Ring Gyro of Arbitrary
Shape and Rotation Axis" American Journal of Physics
Vol 50 No 7 (Jul 1982).
- Michael J Mumma- Laboratory for Extraterrestrial Physics, NASA
Goddard SFC, personal communication (Feb 1987).

E J Post- "Sagnac Effect" Reviews of Modern Physics Vol 39
No 2 (Apr 1967).

Rasor Associates, Inc- Atmospheric Laser System (ATLAS)
Feasibility Investigation RAP-063 (Mar 1981).

Murray Sargent III- "Basic Ring Laser Gyro Theory" SPIE
Vol 487 Physics of Optical Ring Gyros (1984).

CHAPTER 6

PLANETOLOGY AND ASTRODYNAMICS

Chapter Abstract - Engineered planetary lasers at both Mars and Venus are theoretically possible. Because of orbital and gravitational parameters, Venus is the simpler site and therefore preferable. The pentagonal resonator can be deployed in a near-equatorial orbit, and Venus' L1 and L2 libration points provide useful sites for transmitter stations.

In our solar system, natural CO₂ lasers which we might hope to tame exist at both Mars and Venus. The best data and models available [Deming & Mumma, 83] [Gordiyets & Panchenko, 82] indicate a single-pass gain of between 0.05 and 0.10 at the proper altitude for each planet. Thus our reasonable assumption of 0.07 applies as a mean value to both planets, and our choice between them as the best site for this project must hinge on other factors.

The Martian Laser

Except for Mercury and Pluto, Mars has the greatest orbital eccentricity of any known planet. At 0.093377, this measure of the ellipticity of Mars' orbit is over five times that for Earth. A direct consequence is that Mars' aphelion and perihelion distances vary by over $4(10^7)$ km during the 687 d Martian year. Thus the solar "constant" at Mars is about 22 % greater at perihelion than the value normalized to Mars' semimajor axis, and about 17 % lower at aphelion. Variations in the solar pumping rate of the mesospheric inversion follow insolation changes, of course; since Deming & Mumma's gain value is also normalized to semimajor-axis solar distance, we expect single-pass gain to vary between about 0.058 and 0.085, on a seasonal basis, for the subsolar point. Superimposed on this variation will be any additional decrements resulting from orbital motion, as discussed in Chapter 5.

Only two orbits around Mars provide reasonable sites for a pentagonal planetary resonator. Mars is large enough for hydrostatic equilibrium to dominate its shape; since it spins

at about the same angular rate as Earth (24.6 hr), it is subject to substantial polar flattening (ellipticity is 0.009, about three times greater than Earth's). This equatorial bulge provides a noncompensated gravitational attraction which gradually torques an orbit plane around in inertial space. The consequent nodal regression is the most well-known and useful of non-Keplerian orbital perturbations, as it allows establishing an orbit which will turn at the same rate that the planet orbits the sun. Sun-synchronous orbits are commonly used at Earth for remote sensing, since the constant shadow angle on overflown terrain allows direct comparisons among photographs. The Mars Observer mission now planned for 1992 will use a sun-synchronous orbit to get consistent surface images [JPL, 83]. For other applications, a sun-synchronous orbit over the terminator will remain in sunlight, stabilizing both solar power and thermal inputs to a spacecraft. For our purpose, a sun-synchronous orbit can be chosen (Appendix A6-1) which would always overfly the subsolar point throughout the Martian year. Although gravitationally linked to Mars, of course, such an orbit can be thought of mathematically as dependent not on Mars, but rather on the sun; the single-circuit gain it allows varies only with Mars' changing solar distance.

Alternatively, we might choose a Martian equatorial orbit. Since such an orbit has no ascending node, the planetary equatorial bulge has no effect on it. However, Mars' obliquity (defined as the tilt of its rotation axis from the normal to its orbit plane) is, at 25° , comparable to Earth's. That means that an equatorial orbit is subsolar only twice a year, at the vernal and autumnal equinoxes. The departure is greatest at the solstices, where it equals the value of Mars' obliquity. Although we know the cosine law yields only a 10 % decrement for 25° , this variation will compromise further the already-varying seasonal planetary insolation, thus reducing

available gain. The total effect, as numerically modeled by the program MARSYEAR (Appendix A6-2), is shown by Figure 6-1. Because of the unfortunate near-coincidence of Mars' summer solstice with aphelion passage, the minimum single-pass gain value at that time is only about 77 % of the normalized value, or approximately 0.054.

Any orbital inclination greater than zero (equatorial) will fall prey to nodal regression at Mars. Thus while we might make the orbit subsolar in one season, its line of nodes would torque around Mars' rotation axis until at some other time its available gain suffered even more than does that of the equatorial case. The two best orbits are therefore the equatorial and the subsolar sun-synchronous. A direct comparison in Figure 6-2 shows that the advantage of the sun-synchronous orbit is at all times positive-definite. Its maximum is larger, its minimum is less severe, and its variation is more symmetric; by enabling us to escape all penalties save those imposed by Mars' orbit and the fact that satellites must move, such an orbit is the best we can do at Mars. Since the minimum available gain will constrain our communication efficiency, at Mars the value of single-pass gain we could count on would be 0.058.

Of the three major terrestrial planets, Mars is by far the most unlike a sphere. This can be seen qualitatively in Figure 6-3, which compares Mercator-projected surface topography for Venus, Earth and Mars. Mars' small size makes even average topography relatively larger; but the absolute size of features scarring the planet's geologically fascinating face is also awesome. There we find both the largest canyon (Valles Marineris) and the largest mountain (Olympus Mons) known in the solar system. Figure 6-4 identifies Mars' surface topography. Martian macro-topography also varies seasonally, as an amount of CO₂ comprising 10^{-8} the total planetary mass

alternately freezes at each pole [Chao & Rubincam, 87], sublimed and transported atmospherically during Martian spring and fall. The result of this major climatic redistribution is a 25 % annual variability in the gravitational oblateness.

Mars also has the bumpiest gravity field, whose variations do not correlate too obviously with surface features, especially the great differences between northern and southern hemispheres. The largest gravity anomaly "by far" on Mars does correlate with the Tharsis plateau, a region of four great shield volcanoes which includes Olympus Mons [Sjogren, 79]. Tharsis is thought to exist because of an absence of tectonic activity on the planet. Thus, unlike Earth's Hawaiian islands, which evolve sequentially as a crustal plate moves over a stable mantle hot spot, Tharsis remained where it was during its eruptive era, continually piling up material until the planet acquired a substantial bulge. Newton showed that the gravitational force exerted by a homogeneous spherical mass is identical to that exerted by a point of the same mass, and Keplerian orbital motion is based mathematically on that assumption. However, the real gravity field of a non-perfect object will only be spherical infinitely far away, as sketched in Figure 6-5. Satellites orbiting a planet will be tugged around by those ripples in the gravity field; the closer they orbit and the bigger the ripples, the worse the effect.

One disadvantage of the near-polar sun-synchronous orbit at Mars is that it eventually overflies every portion of the planet. Thus whichever planetary gravity anomaly is largest will perturb the resonator orbit, both by accelerating and decelerating the satellites in their orbital path and by causing their altitude to vary. The more severe the tug, both in magnitude and in suddenness, the harder it will be to compensate. As mentioned in Chapter 5, too much variation will cause the line of sight between satellites to miss entirely the

thin inversion layer, resulting in a precipitous and abrupt drop in system single-circuit gain. Gravity anomalies can be modeled in different ways; Esposito et al [86] specify the largest Martian anomalies directly in terms of their local contribution to the gravitational parameter μ . The more standard approach is to represent a planetary gravitational potential in terms of a harmonic expansion in spherical polar coordinates (Appendix A6-3), then determine the coefficients of the model from actual tracking data. Balmino et al [82] have done this to degree and order 18 for most of Mars based on Viking orbiter data. Although the effect of any particular anomaly "is embedded in a series of harmonic coefficients usually not directly identifiable" [Esposito et al, 86], knowing that the J_3 term dominates altitude variations [ESA, 82] enables us to combine Mars Observer predictions with the size of Tharsis to estimate an upper bound on the seriousness of gravitational perturbations for a Mars laser.

Doing this in Appendix A6-4 for the pentagonal case, we conclude that altitude variations of up to 52 km can occur over a time scale of ~18 min, (and smaller amplitudes on a correspondingly shorter time scale) as the resonator satellites orbit the planet. Considering that Deming & Mumma model the inversion layer as only ~10 km thick, a pentagonal Martian resonator would be difficult to keep going. The problem would be exacerbated by irregularities in Mars' atmosphere as well, since the inversion layer can be expected to follow isobaric and isothermal patterns in the mesosphere of which it is a part. Mars' atmosphere is oblate with flattening $(r_e - r_p)/r_e = 0.013$ [Taylor, 76]. This means that the inversion layer can deviate from a circular section by up to 45 km. Additionally, the fluid atmosphere will exhibit peaks and depressions corresponding to gravity anomalies beneath it. It might be thought at first that atmospheric and orbital variations would compensate each other, but the atmospheric

variations affect the line of sight's center, while orbital variations affect its endpoints, and the two are out of phase by π/n where n is the number of resonator satellites. The final effect is unknowable without detailed numerical modeling based on extensive in situ measurements, but its seriousness is evident.

Altitude variations could be compensated propulsively, at a tremendous logistics cost. Efficient motors (like ion engines) are too low-thrust to effect quick maneuvers, and a non-consumable propulsive technique like solar pressing is both too slow and cannot operate during darkside passage (when the polygon continuity must still be maintained). An interesting option would be to use a momentum-storage tether as shown in Figure 6-6. A distant counterweight would establish the satellite center of mass somewhere along the tether; this mathematical center of mass would follow the non-Keplerian mathematical orbit around the planet, popping up and down with the gravity anomalies. By reeling the tether in and out at a rate commensurate with those changes and the changes in gain altitude, the resonator mirror station at one end of the tether could be maintained on a desired path, allowing atmospheric lasing. Energy expended when reeling the tether in could be recovered when paying it out. Magnetic suspension of the mechanism would minimize both vibration transfer to the optical satellite and dissipative friction [Lawing et al, 87], maximizing conservative operation. Tethers for use in space are well characterized theoretically and planned for experimental use almost immediately, so a 100 km tether for orbital station-keeping is not far-fetched. We would expect unusual strength and dynamic difficulties due to the extreme speed with which the Mars laser tether would need to be wound and unwound, though.

At this point it is important to show exactly how much the pentagonal resonator costs at Mars in terms of the stationkeeping just discussed. Appendix A6-5 follows directly the example of Appendix A6-4 to approximate both the altitude variations expected and the maximum system output possible (accounting only approximately for reflection losses) for both the triangular and square cases, which orbit farther from the planet. The extreme attenuation of perturbing gravity effects argues strongly for a triangular geometry at Mars, as does the smoother gain envelope and greater system gain value compared to the square case. However, as noted in Chapter 5, a lossy system gain of 0.03 permits virtually no additional inefficiencies, inaccuracies or uncertainties (all of which are sure to be present in a system of this scope) before no circulating energy is left to couple out. This grand tradeoff between system circuit gain and circuit continuity shows that, although probably feasible, engineering a planetary laser at Mars would seem a marginal undertaking of great risk, requiring extremely cautious and detailed study.

Other operational system implications at Mars derive from its location in the solar system and its place in long-range human planning. Solar energy for spacecraft housekeeping is less available, but thermal rejection is also easier, given Mars' distance from the sun, than at any other terrestrial planet. Mars has two tiny moons, Phobos and Deimos; although they have such little mass that each is generally ignored when calculating the other's orbit, they nonetheless would produce perturbations, particularly the inner moon Phobos, which has a periapsis altitude of only 8811 km. On the other hand, Phobos provides a ready platform for maintenance operations and a source of materials for construction and life support.

Extensive human activity is planned for the Martian neighborhood and the planetary surface itself. While this

means that the entire infrastructure necessary to build a planetary laser would have a head start here, and that system economies would result from substantial ongoing overlap between this and other projects, it also means that the Martian orbital vacuum environment will inevitably become materially degraded. For optical systems in general this is regrettable, but for an elaborate optical system with enormous mirrors (like a planetary laser) operating on the edge of feasibility, contamination would be expensive at best and disastrous at worst. Given that nuclear-electric propulsion will be a staple freight technology in future solar system development, it is fortunate that inert gases make the most useful fuels for ion engines [Aston, 87]. However, where there are people there will be oxygen (leaks), which can be photoionized by solar energy, and even low fluxes of atomic oxygen have already proved to reduce severely orbital material lifetimes [Whitaker et al, 87].

These more subtle final issues are overwhelmed by the seriousness of the physics outlined above. While it seems clear that if Mars were the only planet in our solar system with a CO₂ atmosphere capable of supporting natural lasing, we could arrange stunts to use it, the truth is that the Red Planet is not an inviting environment for engineering planetary lasers. But with serendipity perhaps second only to the existence of Earth and life itself, our solar system provides us with the choice of another planet sustaining a natural laser.

The Venusian Laser

The only human plan for Venus discussed to date is eventual terraforming; the planet in its present form is almost completely unsuitable even for the most mechanically supported human life imaginable. With an atmosphere 90 times denser at the surface than Earth's, a surface temperature high enough to melt zinc, permanently dense sulfuric acid clouds which only let 3 % of the sun's light through to the surface, and a gravity well virtually as deep as Earth's, Venus and its orbital neighborhood are sure to remain intact and unpeopled until terraforming becomes possible. (The least of the requirements would be stationing a vast sunscreen at the L1 point to reduce insolation and reverse planetary greenhouse warming; then massive amounts of hydrogen would have to be added, probably by purposeful asteroidal bombardment.) Venus is as interesting a planet as Mars, if less spectacular. It has no natural satellites, and virtually no intrinsic magnetic field. Its orbit is the most circular of any major inner planet. It is hardly tilted, barely rotates, and is about as close to a perfect sphere as a real planet could be. Ironically, our "twin" planet and closest neighbor in the solar system, inimical though it is to human life, seems to provide an almost perfect site for operating an informational link to other star systems.

Venus' orbit, having eccentricity 0.006787, contrasts greatly with Mars' --- it is 14 times less elliptical, and 2.5 times less elliptical than Earth's orbit. In fact, the difference between its perihelion and aphelion solar distances is only $1.5(10^6)$ km, or just over 1 % of its semimajor axis distance. Thus the total insolation change over the course of a 224.7 d Venusian year is just 2.7 % (remember that Mars' total variation was about 39 %). Since the decrement from average suffered at aphelion is therefore only about 1 %, we

will neglect it for the rest of this work, considering the seasonal insolation at Venus to be constant.

Venus rotates extremely slowly, one retrograde sidereal turn taking 243 d (longer than its year!). Consequently its polar flattening is only about 10^{-5} . Furthermore, its figure - mass center offset is calculated to be a mere 430 m (contrasting with kilometer-scale offsets for Mars and Earth), so that the "most striking feature" of the planet's shape is its "extreme sphericity" [Pettengill et al, 80]. Thus the J_2 coefficient in Venus' gravity field expansion is vanishingly small --- of order $5(10^{-6})$ [Williams & Mottinger, 83] --- implying that insufficient oblateness exists to cause nodal regression of non-equatorial orbits [Uphoff, 79]. No regular sun-synchronous orbits are possible at Venus. However, the absence of nodal regression also means that a circular orbit in the plane of Venus' orbit around the sun will remain in that orientation essentially forever, regardless of its relation to Venus' equator. (The equatorial obliquity is a mere -2.6° anyway.) By choosing a near-equatorial resonator orbit at Venus, then, we avoid the astrodynamical limitations found at Mars; the satellite orbit always passes over the subsolar point, and first order resonator output varies only according to the pentagonal ripple discussed in Chapter 5.

All that remains is to investigate what kind of planet Venus is to orbit, since Mars proved that gravity variations can be a handicap. First we note that even small anomalies will perturb long-term satellite motion because Venus' slow rotation insures that many consecutive orbits will see the same terrain at about the same true anomaly [Uphoff, 79] [Mohan & Esposito, 84]. The Magellan Venus mission, whose elliptical radar mapping orbit ranges in altitude from 300 km to 7760 km, expects a 13 km periapsis altitude increase over the 243 d nominal mission lifetime due both to long-term gravity

effects and solar third-body perturbation [Cutting et al, 84] [Kwok, 84], which of course is substantial at Venus. These predictions, based on a 10th-degree and -order spherical harmonic model of Venus' gravity calculated from Pioneer Orbiter Differential Very Long Baseline Interferometry Doppler tracking data, are those actually used to plan the Magellan mission profile. The $\Delta v = 20$ m/s propellant allocation required over the mission life for orbit trim maneuvers to correct such long-term perturbations is easily within an annual logistics budget for a planetary laser, as shown in Chapter 4.

As at Mars, where they pose a severe operational problem, short-term orbit altitude variations at Venus depend on the bumpiness of the planetary gravitational potential. Fortunately at Venus the pentagonal resonator satellites orbit about twice as far from the surface as they would at Mars, and Venus' much greater planetary radius means that anomalies will not in general distort the gravity field as much as at the smaller planet. Indeed, variations modeled by harmonic coefficients above degree and order 7 have virtually no effect on orbits over 1400 km high at Venus [Williams & Mottinger, 83], a great benefit since gravity events with higher spatial frequency require quicker compensation. The pentagonal orbit altitude is 1589 km.

At Venus, the "amplitudes of the gravity anomalies are not at all like those for...Mars", being instead "relatively mild" with "amplitudes similar to those on the earth" [Sjogren et al, 80]. Unlike Mars and the Earth though, the anomalies correlate well with topography [Esposito et al, 82]; however, both Venusian "continents" are mostly isostatically compensated, as on Earth and the Moon [Masursky et al, 80]. This means that these features consist of material less dense than the lithic layer they "float" on; consequently their reduced relative mass

cancels the effect on their gravity signature of their height. Aside from these isolated features "the planet is quite flat" [Pettengill et al, 80]. Note again Figure 6-3 to see qualitatively the tremendous topographical distinction of Venus from the other terrestrial planets. From Pioneer radar altimetry we have the most complete, detailed set of global altitude data for Venus than for any planet except ours [Masursky et al, 80], and they show that despite a 13 km total topographical relief, only 8 % of the mapped surface exceeds its mean altitude by more than 1.5 km.

Figure 6-7 compares hypsometric distributions (differential topographic spectra) for the Moon, Mars, Earth and Venus. Mars' trimodal distribution shows clearly the distinctions among the Tharsis plateau, southern highlands and northern lowlands. The Moon's weakly bimodal graph reflects its distinct highlands and maria, while Earth's strongly bimodal distribution shows clearly that 30 % of its surface comprises continental plateaus above ocean basins. Venus' graph is unique both by being unimodal and by having its lone peak so narrow [Masursky et al, 80]. Fully 60 % of Venus' mapped surface is within 500 m of the modal planetary radius, and 20 % is within 125 m [Pettengill et al, 80].

Figure 6-8 locates Venus' major geological features. Beta Regio is probably one of the youngest regions on the planet, consisting of two volcanic shields [Masursky et al, 80]. Correlated with Beta is the planet's largest gravity signature [Esposito et al, 82]. Ishtar Terra is the northern hemisphere highland region, and contains Maxwell Montes, Venus' highest mountain. Maxwell exhibits no gravity anomaly, and the rest of Ishtar appears mostly compensated also, as indicated earlier. The largest highland on Venus is Aphrodite Terra, a region about the size of South America located across Venus' equator. Clearly the worst short-term altitude variations experienced by

near-equatorial satellites will be caused by Aphrodite, although we expect these to be small since its gravity signature is largely isostatically compensated as noted [Masursky et al, 80]. Appendix A6-6 calculates that we would expect no worse than 2 km altitude variations for the pentagonal resonator at Venus. Given a gain layer about 10 km thick, this orbital condition means roughly that for mirror diameters less than 6 km, pointing and tracking, not gravitational effects on the orbit, will limit the system. We can consider satellite altitude variations and their rate to be below a "disturbance threshold" for resonator continuity.

Venus does provide one inconvenience, however. While we might presume that its mesospheric altitude varies as little as the planetary gravity field which shapes it, a much larger effect results from inherent diurnal variations [Ananda, 80]. Insolation during Venus' long days causes atmospheric swelling, so that density varies erratically (as much as an order of magnitude at the same altitude) near the terminator, and increases appreciably toward noon [Sjogren et al, 80]. Although the difference between dayside and nightside is irrelevant to us, the fact that the dayside inversion layer probably does not form a spherical shell means that single-circuit gain cannot be modeled as simply as we have done. Figure 6-9 compares the effect on the available gain envelope of eliminating various amounts from the beginning and end of each sector's contribution as a way of discounting its unreliable performance near the terminator. Until the Venusian inversion layer is mapped in fine detail, we cannot know just how much of the terminator region to discount, but Figure 6-9 shows that the effect could range from striking to severe. Should it turn out that substantial gain altitude variations do follow the atmosphere's diurnal "breathing", then a higher order resonator polygon (providing more sector overlap) might be required, or a non-circular orbit (forming a non-regular

pentagon) might compensate. In the worst case, we might resort to a tether mechanism for moving the satellites, as proposed for Mars. For this study we will assume that a circular orbit works for the pentagonal case.

Given a steady-state ring resonator around Venus, we require some method to couple laser light out of the moving system, modulate it and direct it to a receiver target somewhere on the celestial sphere. Most desirable would be to station a relay satellite at some position fixed with respect to the Venus system, so that only the orbital motion of the laser coupler would have to be compensated. As mentioned in Chapter 5, such points in space do exist: the five Lagrange libration points. First studied by Lagrange, they are the only points in a rotating system defined gravitationally by two primary masses, at which a third, much smaller, mass will experience no relative acceleration. At these locations, the attractive force for both primaries exactly equals the product of the satellite's mass and centripetal acceleration due to its motion about the system barycenter. Figure 6-10 diagrams the arrangement. The two points designated L4 and L5 are referred to as the triangular points because they lie on the smaller primary's orbit, 60° ahead of and behind it. These are stable positions, because an object in either vicinity will move toward that point; the regions are known to collect debris as demonstrated by the Trojan asteroids in Jupiter's orbit. The other three, collinear, points lie on the line connecting the primaries. L1 lies between the two primaries; L2 lies beyond the small primary's farside; L3 lies beyond the large primary's farside. These three are unstable positions since an object in any of their vicinities will move away from them unless either stationed precisely, orbited about the libration point in a "halo" orbit (as demonstrated by the ISEE-3 spacecraft), or maintained with a small propellant budget.

Each pair of bodies in our solar system produces a libration set: the Earth - Moon system, the Earth/Moon - Sun system, and so on including the Venus - Sun system. Although calculating the collinear point positions analytically is not trivial, tables of these values for the major solar system bodies may be found in [CSC, 77]. Changing the primary separation will of course move the libration points; here again Venus' low orbital eccentricity benefits the stability of planetary laser operation. For locating our relay station, L3 is useless for us since it orbits behind the sun. L4 and L5 are separated from Venus by its solar distance; such an interplanetary-scale line of sight, while it reduces the need to slew an optical relay which tracks Venus' orbiting resonator satellites, can only exacerbate system pointing accuracy requirements. L1 and L2, however, are only about 10^6 km away from Venus and situated diametrically. A mirror system located this far away would still only have to tilt less than 10° to maintain a constant incidence angle for a beam coupled out from any point along the resonator satellites' orbit. As outlined in the next chapter, establishing two identical relay modulators at Venus' L1 and L2 points will enable us to avoid incidence angles worse than 45° while maintaining continuous transmission to virtually any point on the celestial sphere over times ranging from minutes to years.

Of the two available planets known to support natural CO₂ lasers, Venus provides the most attractive site for engineering a planetary ring resonator. Local orbital mechanics does not preclude establishing a pentagonal resonator configuration stable over long times, and also permits locating beam-modulating relays at unmoving positions along the Sun - Venus line, to facilitate the large duty factors desirable for interstellar communication links. We therefore choose Venus for the planetary laser site. Astrodynamical values pertinent to this case are collected for easy reference in Appendix A6-7.

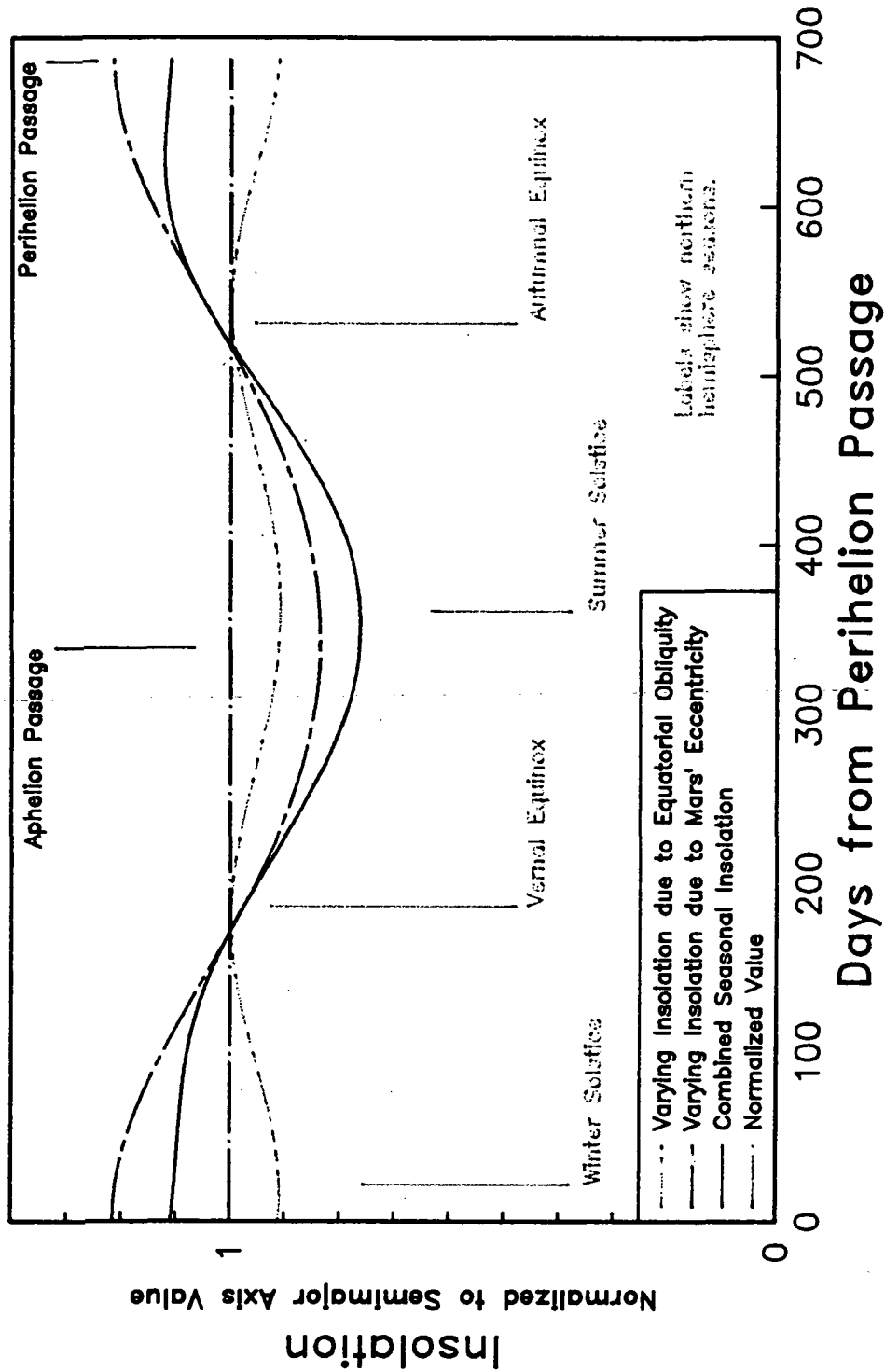


Figure 6-1 Seasonally varying insolation available beneath an equatorial orbit at Mars.

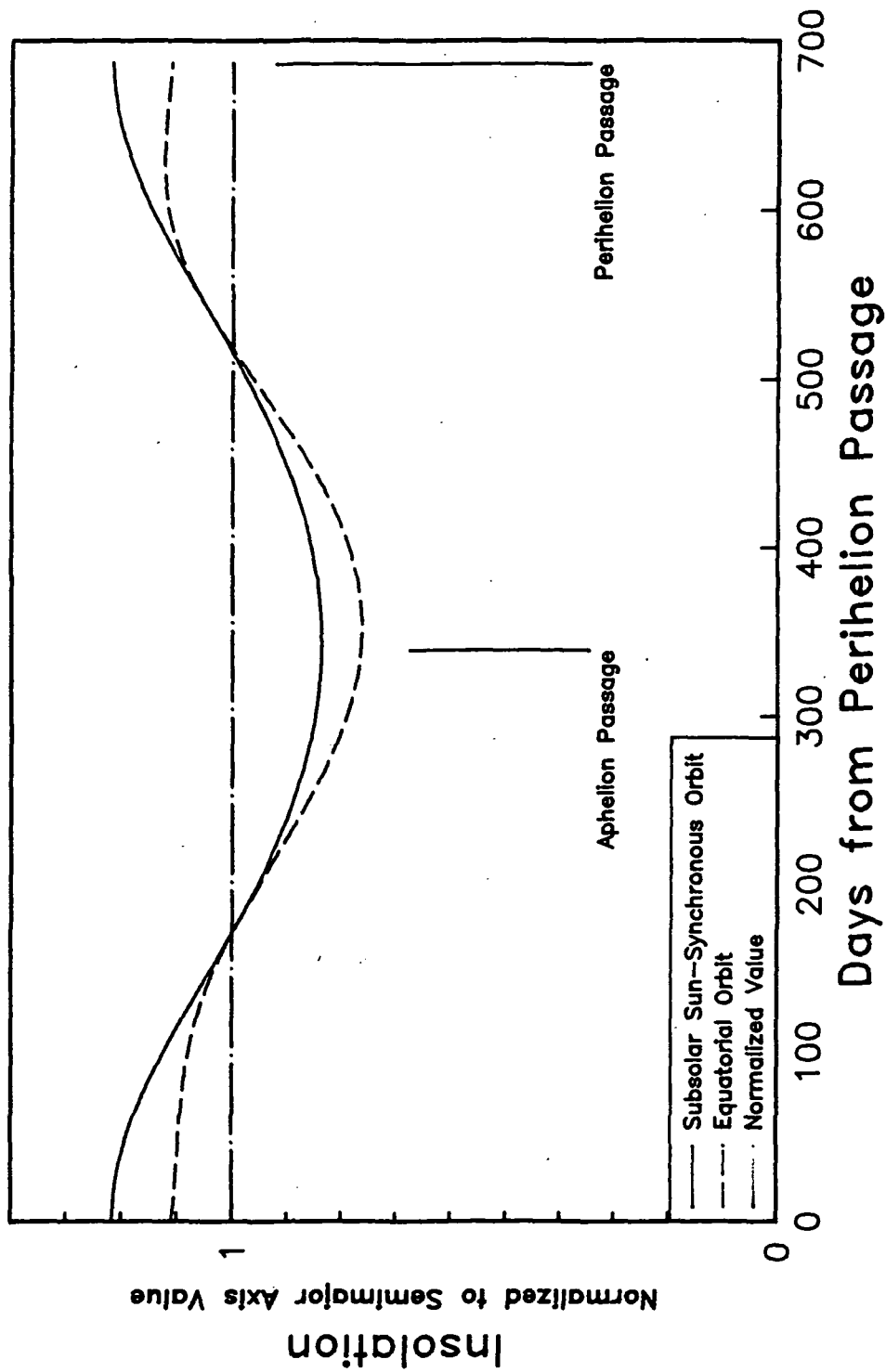


Figure 6-2 Comparison of the seasonal insolation available beneath equatorial and subsolar sun-synchronous orbits at Mars.

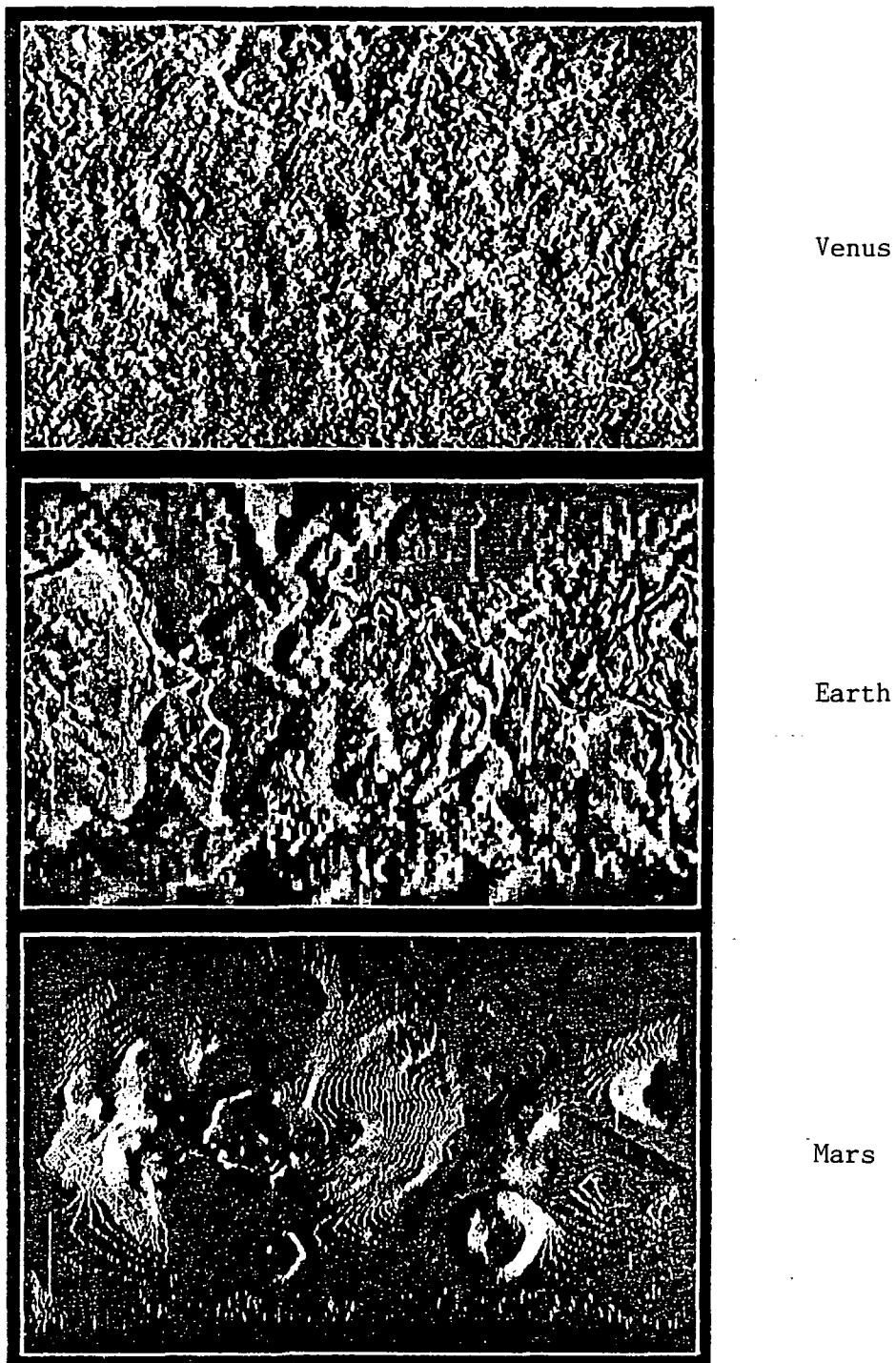


Figure 6-3 Mercator-projected topographical maps of the major terrestrial planets, generated from radar altimetry data and adjusted to the same area [Hartmann, 83].

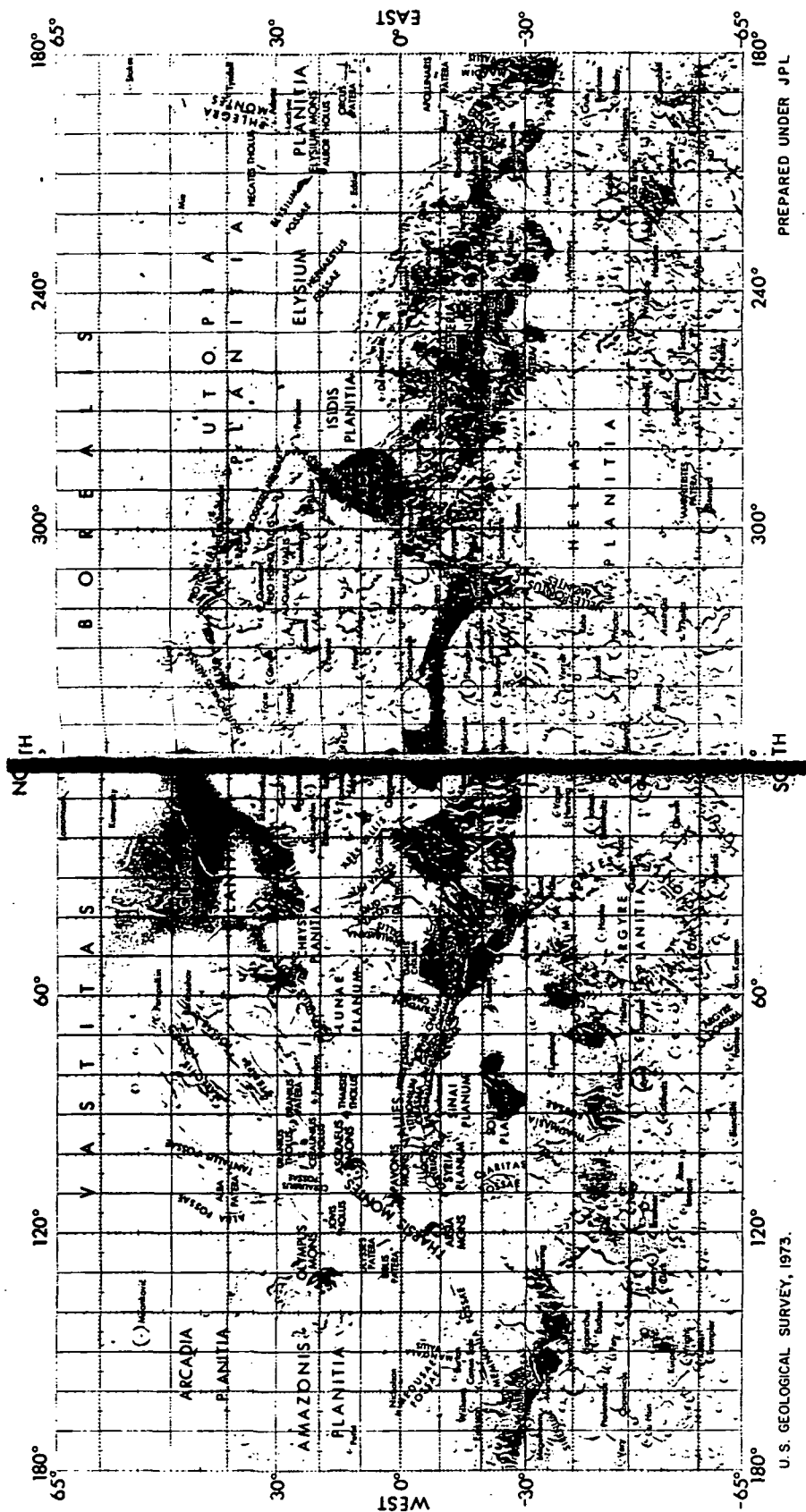


Figure 6-4
Geographical
map of Mars.
[Blunck, 82]

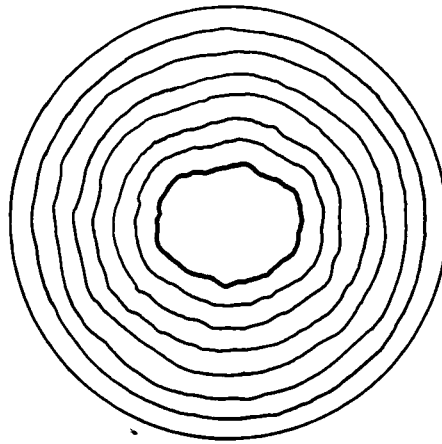


Figure 6-5 Equipotential surfaces of a planetary gravity field, showing how the effects of non-sphericity diminish with increasing radial distance. [Vanicek & Krakiwsky, 82]

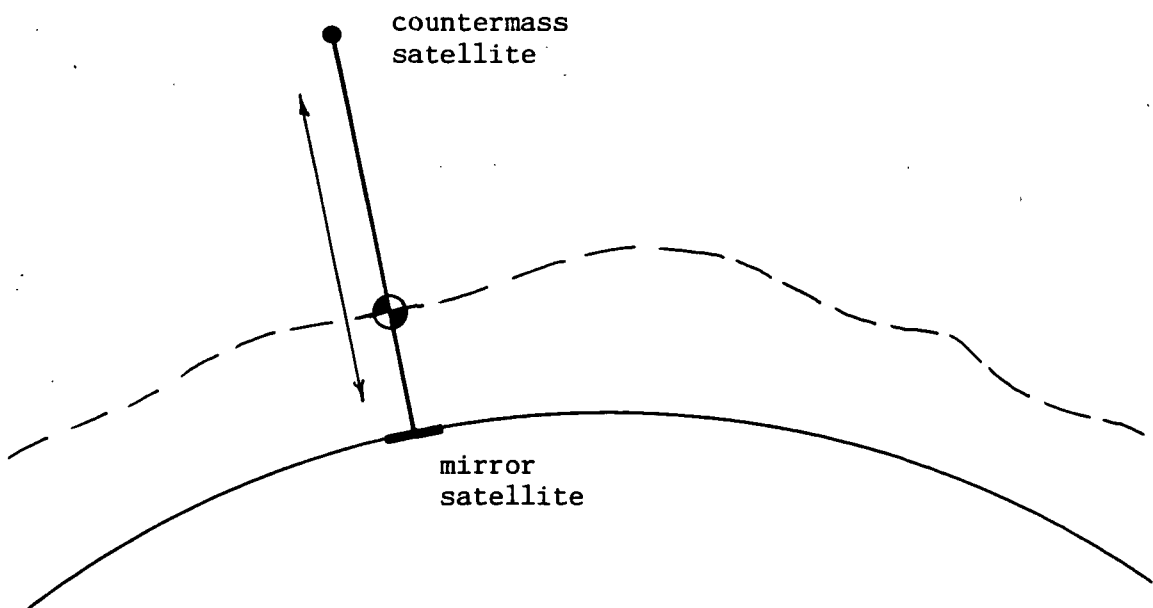


Figure 6-6 A scheme for counteracting short-term orbit altitude perturbations. The center of mass will at all times follow the real, non-Keplerian orbit, but by compensating with active tether length changes, the critical mirror satellite could be made to follow a prescribed path.

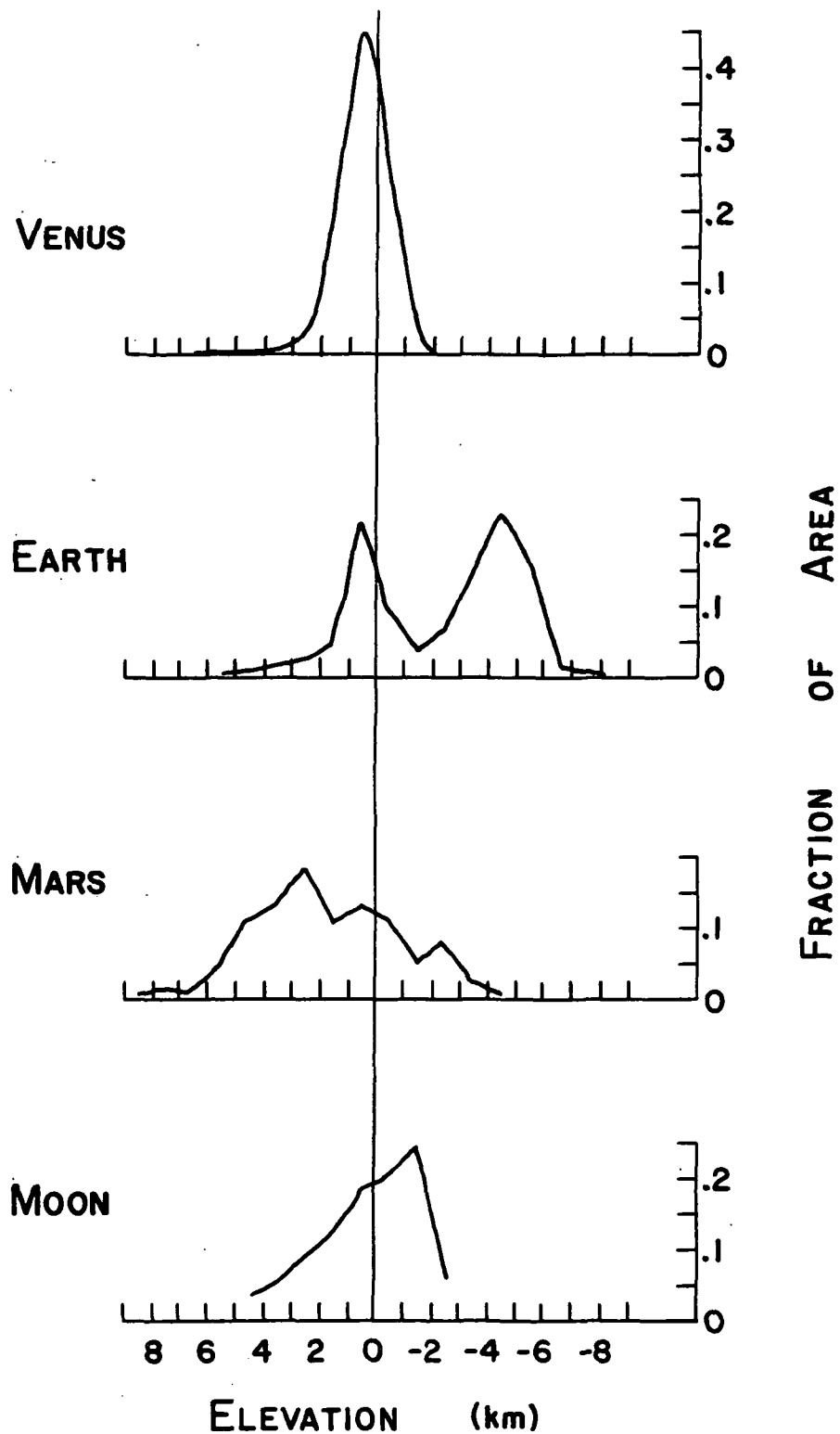


Figure 6-7 Comparative hypsometric distributions for the major terrestrial planets around their mean radii, adjusted to a normalized area. Adapted from [Masursky et al, 80]

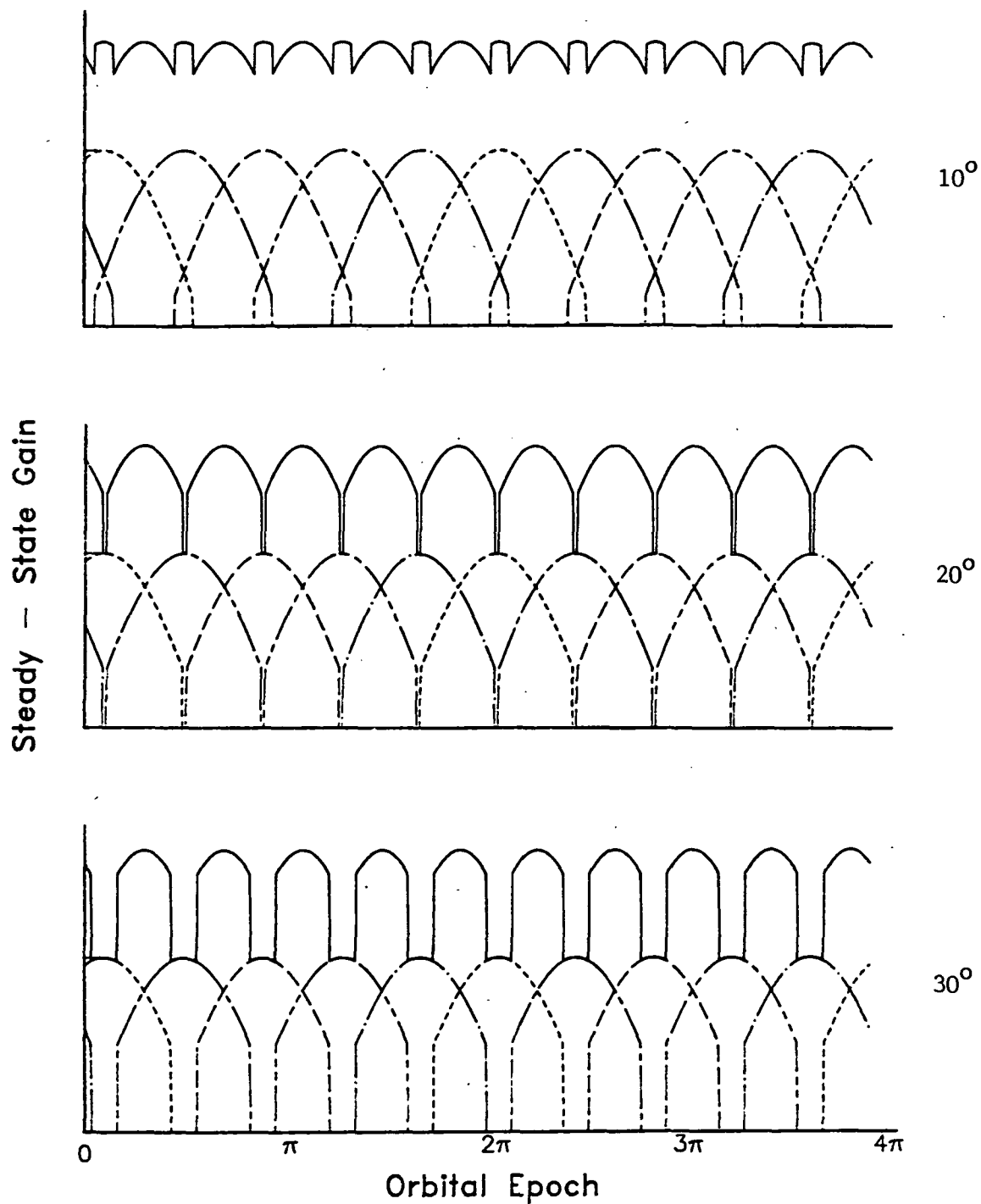


Figure 6-9 Available single-circuit gain envelopes for the pentagonal resonator at Venus after discounting the indicated amount from the beginning and end of each sector's dayside pass, to model lasing unreliability of the terminator region.

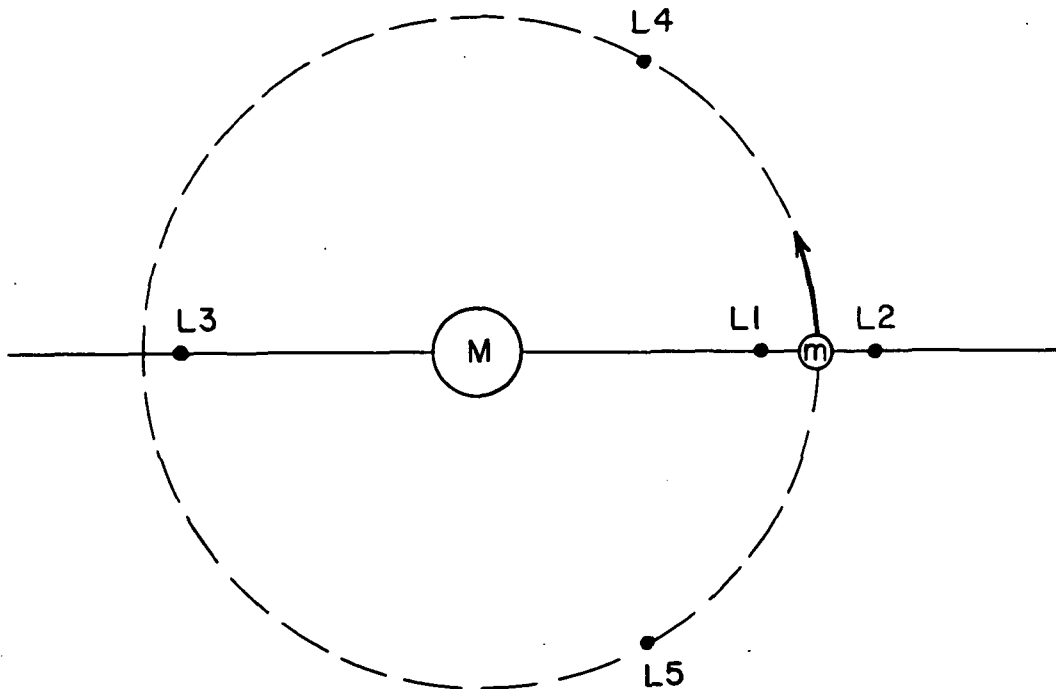


Figure 6-10 Diagram (not to scale) showing convention for labelling the Lagrange libration points inherent in a rotating system comprised of primary mass M and secondary m .

Appendix A6-1 Sun-synchronous orbits at Mars.

The basic equation describing nodal regression is [Agrawal, 86]:

$$\dot{\Omega} = -\frac{3}{2} \frac{\mu}{h} \frac{J_2}{r^3} R^2 \cos i \quad (\text{A6-1.1})$$

where $\dot{\Omega} \equiv$ the time rate of change in longitude of the orbit's ascending node, $h \equiv$ the scalar value of orbital angular momentum, $\mu \equiv$ the planetary gravitational parameter (GM), $J_2 \equiv$ the 2nd zonal coefficient in the spherical harmonic expansion of the planetary gravity field (Appendix A6-3), $r \equiv$ the orbital radius, $R \equiv$ the planetary radius, and $i \equiv$ the orbital inclination.

Designing a sun-synchronous orbit consists of choosing the desired $\dot{\Omega}$ and r for some planetary mission, and calculating the orbital inclination i required to make it work:

$$i = \cos^{-1} \left[-\frac{2}{3} \frac{\dot{\Omega}}{J_2 R^2} \left(\frac{hr^3}{\mu} \right) \right] \quad (\text{A6-1.2})$$

This can be simplified for our case of circular orbits. Since the radius and velocity vectors of a circular orbit are orthogonal, the definition $h \equiv r \times v$ yields the scalar identity $h = rv$.

Furthermore, the constant tangential speed in a circular orbit is:

$$v_{cs} = \left(\frac{\mu}{r} \right)^{\frac{1}{2}}$$

Thus the last term in equation A6-1.2 can be reduced:

$$\frac{hr^3}{\mu} = \frac{r^4}{\mu} \left(\frac{\mu}{r} \right)^{\frac{1}{2}} = \left(\frac{r^7}{\mu} \right)^{\frac{1}{2}}$$

Then equation A6-1.2 becomes:

$$i = \cos^{-1} \left[-\frac{2}{3} \frac{\dot{\Omega}}{J_2 R^2} \left(\frac{r^7}{\mu} \right)^{\frac{1}{2}} \right] \quad (\text{A6-1.3})$$

The orbit should regress once per Martian year. Hence:

$$\dot{\Omega} = + \frac{2\pi}{687(24)3600} = 1.06(10^{-7}) \frac{\text{rad}}{\text{s}}$$

In addition, substitute standard Martian parameters:

$$J_2 = 1.96(10^{-3})$$

$$R = 3398 \text{ km}$$

$$\mu_{\sigma} = 4.305(10^4) \frac{\text{km}^3}{\text{s}^2}$$

The orbital radius for our case is specified by the resonator geometry chosen:

$$r = \frac{R + h}{\cos \gamma} \quad \text{where } \gamma = \frac{360^\circ}{2n}$$

if $h \equiv$ the altitude of the mesospheric inversion layer and $n \equiv$ the number of sides in the polygon (number of lasing sectors). At Mars, $h = 70 \text{ km}$, and the inclinations of sun-synchronous orbits corresponding to several resonator geometries are given below:

<u>geometry</u>	<u>orbital radius</u>	<u>orbital inclination</u>
triangular	6936 km	114.7°
square	4904 km	97.1°
pentagonal	4287 km	94.5°

Typical of sun-synchronous orbits, these are all retrograde ($i > 90^\circ$), and the last two are considered near-polar. If these orbits are established such that they cross Martian noon, they will stay that way throughout the Martian year to first order.

Appendix A6-2 Program MARSYEAR.FTN for calculating the insolation available to a planetary resonator in Martian equatorial orbit.

The program incorporates those effects of Mars' complicating orbital eccentricity which can be simply modelled mathematically. Although the core of the program is simply four equations, the subtle and important assumptions behind those equations are explained on the following pages.

```
ftn,1,s
$files 0,1
      program marsyear
      real nom
      open(99,file = 'year.dat')
      pi = 4*atan(1.)
      estep = 2.*pi/687.
      e = 0.
      nom = 1.

2 solecc = (1./(1. - .093377 * cos(e)))**2

      truanm = 2*atan(1.098175 * tan(e/2.))
      solobl = .955 - .045*(cos(2.*truanm - .3304))

      sol = solecc * solobl

      d = e/estep
      write(99,*) e, d, nom, solecc, solobl, sol, truanm
      e = e + estep
      if (e.le.2.*pi) go to 2

      close(99)
      stop
      end
```

Two separate effects must be included. First, the suborbital insolation of an equatorial orbit at Mars will vary seasonally since the planet is tipped; only at the equinoxes will the equator be subsolar. Second, the amount of light reaching Mars varies seasonally because of the planet's orbital eccentricity. These two periodic variations combine to limit the potentially available laser gain.

We model these effects by organizing the calculation as two separate functions (solobl and solecc), both of which vary from a normalized insolation value defined as that available at the subsolar point when Mars is at semimajor axis distance from the sun. The combined effect is then easily obtained by multiplying the values of these two functions at each time step.

For a closed orbit, equal time steps mean equal steps in M , the mean anomaly. Whereas these correspond to equal angular increments for a circular orbit, Kepler's 2nd Law precludes this correspondence for elliptical orbits. Although Mars' eccentricity is large for a planet, at less than 10 % it is small mathematically; consequently to keep regular divisions on our abscissa, we will increment eccentric anomaly E as though it were M . Only a slight distortion in time will result. Our step is $2\pi/687$, or one day, and the program runs for one Martian year.

Eccentricity Effect

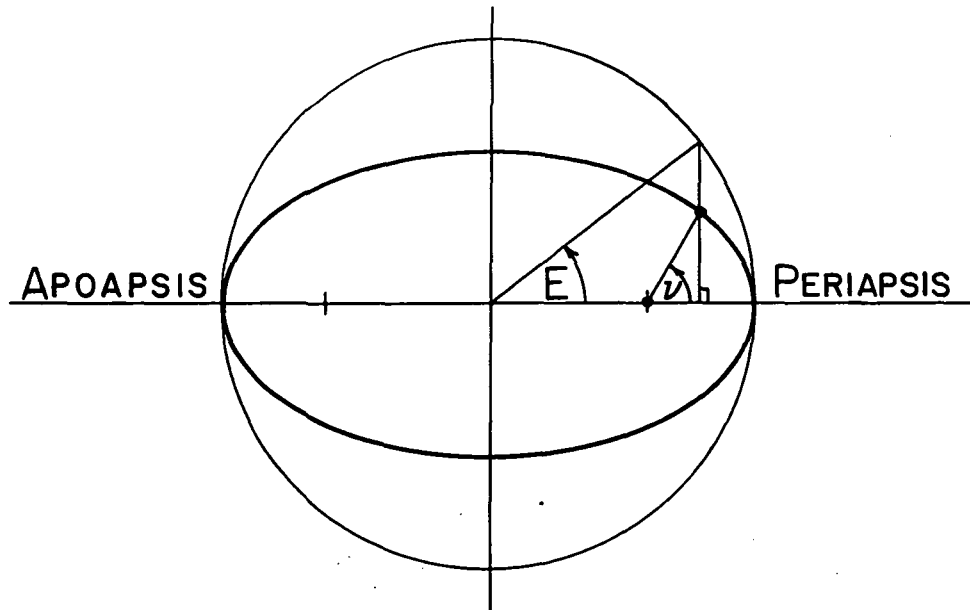
Radiation decreases directly with the square of distance:

$$l \cdot a^2 = I_e \cdot r^2 \quad (A6-2.1)$$

where $a \equiv$ semimajor axis distance, $r \equiv$ length of the radius vector from Mars to the sun at any time, $l \equiv$ normalized insolation, and $I_e \equiv$ insolation as a function of eccentricity. Now for an elliptical orbit:

$$r = a(1 - e \cos E) \quad (A6-2.2)$$

where $e \equiv$ eccentricity and $E \equiv$ eccentric anomaly at any time, which means the central angle swept out along a circle (circumscribing the elliptical orbit) by a point joined to the orbiting object by a perpendicular dropped to the line of apsides:



Substituting equation A6-2.2 into equation A6-2.1, we can solve for I_e :

$$I_e = \frac{1}{(1 - e \cos E)^2} \quad (\text{A6-2.3})$$

which is normalized as it should be by being independent of a .

Allowing E to step forward 2π around the circle, the varying I_e is easily calculated.

Obliquity Effect

Modelling this is more complicated, because the function's maxima must occur at the equinoxes (when the equator is subsolar). Thus the geometry which determines their occurrence centers not on the eccentric anomaly circle we are using for our time step, but rather on the sun, at the primary focus of Mars' elliptical orbit. We must therefore

transform the time step into a measure of true anomaly ν , which can be done at all times through Gauss' Equation:

$$\tan\left(\frac{\nu}{2}\right) = \left(\frac{1+e}{1-e}\right)^{\frac{1}{2}} \tan\left(\frac{E}{2}\right) \quad (\text{A6-2.4})$$

Transforming the regular time steps in this way allows us to build a simple periodic function depending on ν , such that its minima and maxima depend as they should on seasonal cardinal directions centered on the sun. The normalized obliquity-insolation function can take the general form:

$$I_o = K + A(\cos(\omega\nu) + \phi) \quad (\text{A6-2.5})$$

Setting $\omega = 2$ gives two maxima per year as required (there are two equinoxes). Setting $A = .045$ gives a double amplitude (total decrement from the normalized value) of .09, corresponding to the worst case ($\cos 25^\circ$) at solstices. Setting $K = .955$ allows the peak to reach the normalized value of 1.0. Finally, the combination of the sign of A and the value of the phase angle ϕ enables starting this seasonal function at the right time of year.

Mars passes through perihelion about a month before the start of southern hemisphere summer [Hartmann, 83], which is the same as the northern hemisphere winter solstice. Taking one month as 30 d, this interval is a change in mean anomaly of:

$$\Delta M = 2\pi \left(\frac{30}{687} \right) = .2744 \text{ rad}$$

To be as accurate as possible, let us recast this interval in terms of our function's "numerical time", the true anomaly ν . Since all the orbital angular measures start from periapsis crossing, all we need is to find the true anomaly corresponding to $M = .2744 \text{ rad}$.

M is related to E through Kepler's Equation:

$$M = E - e \sin E \quad (A6-2.6)$$

which of course cannot be solved analytically. A numerical solution for a particular M is, however, straightforward.

SEASON	The simple program SEASON, listed here, was run
0	on the HP-41CV programmable calculator to solve
STO 01	Kepler's Equation. It uses a numerical step
LBL 01	for E of $2\pi/400$.
ENTER	
PSE	The results:
RCL 01	E
SIN	M
.093377	.2985
*	.3142
-	.2852
PSE	can be interpolated for M = .2744 to yield
RCL 01	E = .3013, with which Gauss' Equation then yields:
2	
ENTER	
π	$v = .3304$
*	
400	
\div	
+	
STO 01	
GTO 01	
STOP	

With this true anomaly as a phase lag, a negative sign on A will insure that the obliquity function's first minimum (winter solstice) will follow the eccentricity function's maximum (perihelion passage) by about a month.

Thus the particular form of equation A6-2.5 which will effectively model insolation variation due just to Mars' equatorial obliquity is:

$$I_o = .955 - .045 \left[\cos(2v - .3304) \right] \quad (A6-2.7)$$

Appendix A6-3 Conventional planetary gravity model.

This analysis intends to outline the origin of the governing differential equation for planetary gravitational potential, and its solution (for the boundary value problem of interest) into an infinite series representation whose effects can be usefully dissected. For a complete discussion, see [Vanicek & Krakiwsky, 82], [Wertz, 84] and [Kaplan, 76].

Let the gravitational acceleration field be $\mathbf{g}(\mathbf{r})$, a vector function of position. Then the gravitational flux through a differential surface dS is $\mathbf{g}(\mathbf{r}) \cdot \mathbf{n} dS$, where $\mathbf{n} \equiv$ the unit normal to dS . Taking the convention that \mathbf{n} is positive outward, the divergence theorem can be written in the limiting case as:

$$\nabla \cdot \mathbf{g}(\mathbf{r}) = \lim_{V \rightarrow 0} \frac{\oint_S \mathbf{g}(\mathbf{r}) \cdot \mathbf{n} dS}{V} \quad (\text{A6-3.1})$$

where V is a volume enclosed by the surface S .

We seek an analytical expression for the gravitational flux. First, represent the mass M of volume V by $M = V\sigma(\mathbf{r})$, where the density σ is a function of position, but assumed constant over small V . Now taking the coordinate origin at the center of V , Newton's law of universal gravitation gives:

$$\mathbf{g}(\mathbf{r}) = \frac{GM}{r^3} \mathbf{r} \quad (\text{A6-3.2})$$

where $G \equiv$ the universal gravitation constant, and \mathbf{r} is the vector separation from the center of V to the point at which \mathbf{g} is measured. Since the surface integral of equation A6-3.1 is closed, the shape of S is irrelevant and we choose a sphere for simplicity. Then \mathbf{g} is normal to S everywhere, so:

$$\mathbf{g} \cdot \mathbf{n} dS = g dS = - \frac{GM}{r^2} dS = - \frac{G\sigma(r)V}{r^2} dS \quad (\text{A6-3.3})$$

where the minus sign arises because the inward (attractive) acceleration opposes our positive-outward convention for \mathbf{n} .

A sphere of radius r has area $4\pi r^2$, so from equations A6-3.1 and A6-3.3:

$$\nabla \cdot \mathbf{g}(r) = \lim_{V \rightarrow 0} - \frac{4\pi r^2 G\sigma(r)V}{r^2 V} = - 4\pi G\sigma(r) \quad (\text{A6-3.4})$$

Gravity is an irrotational, or conservative, force, meaning that no work is performed in moving a mass around any closed path under its influence:

$$\oint_C \mathbf{F}_g \cdot d\mathbf{r} = 0 \quad (\text{A6-3.5})$$

Now any conservative force field can be represented as the gradient of a scalar potential function, convenient because its value at any point is completely specified by a single number rather than a triplet. In this case:

$$\mathbf{g} = \frac{\mathbf{F}}{m} = \nabla U \quad (\text{A6-3.6})$$

where we define U as the gravity potential. It is this quantity in particular which we want to model as a way of describing the gravity field of a planet. Substituting equation A6-3.6 into A6-3.4, we get:

$$\nabla \cdot \nabla U \equiv \nabla^2 U = - 4\pi G\sigma(r) \quad (\text{A6-3.7})$$

This is the partial differential equation governing the gravity potential produced by a mass at any point not momentum-bound to the mass.

For the region of our interest, outside the flux source M , density $\sigma = 0$ and this Poisson equation becomes the homogeneous Laplace equation:

$$\nabla^2 U = 0 \quad (\text{A6-3.8})$$

known as the gravity potential equation in free space. Solutions to the Laplace equation are called harmonic.

In the spherical polar coordinate system appropriate for describing planets and astrodynamics, equation A6-3.8 takes the form:

$$\frac{\partial^2 U}{\partial r^2} + \frac{2}{r} \frac{\partial U}{\partial r} + \frac{1}{r^2} \frac{\partial^2 U}{\partial \theta^2} + \frac{\cot \theta}{r^2} \frac{\partial U}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 U}{\partial \phi^2} = 0 \quad (\text{A6-3.9})$$

where $r \equiv$ the radial coordinate, $\theta \equiv$ the latitudinal coordinate ($0 < \theta < \pi$), and $\phi \equiv$ the longitudinal coordinate ($0 < \phi < 2\pi$).

Classical solution by separation of variables presumes that the function of three variables can be factored into a product of three single-variable functions:

$$U(r, \theta, \phi) = R(r) \Theta(\theta) \Phi(\phi) \quad (\text{A6-3.10})$$

Since we are interested in modelling the gravitational potential outside the planet, we take as the lower limit of the boundary value problem a "Brillouin" sphere which encloses all the planet's mass, and outside of which Laplace's equation applies. On the surface of any such sphere of radius r , the function $R(r)$ becomes constant and constitutes the solid spherical harmonics. Then the product $\Theta\Phi$, the surface spherical harmonics, accounts for all the variability of U .

For concise algebraic details of the variable separation, see Wertz [84]. Developing the eigenfunctions of the Laplace equation in spherical polar coordinates yields the expanded form of equation A6-3.10:

$$U(r, \theta, \phi) = \sum_{n=0}^{\infty} \left(\frac{a}{r} \right)^{n+1} \sum_{m=0}^n (C_{nm} \cos(m\phi) + S_{nm} \sin(m\phi)) P_{nm}(\cos\theta) \quad (\text{A6-3.11})$$

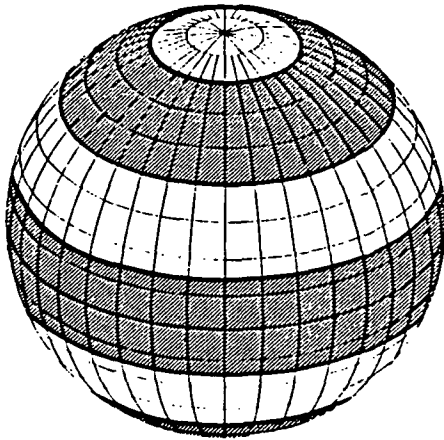
where $n = 0, 1, 2, \dots$, $a \equiv$ the planetary radius, $r \equiv$ the radius of the Brillouin sphere, the C_{nm} and S_{nm} are lists of numerical coefficients, and $P_{nm}(\cos\theta)$ are associated Legendre functions in $\cos\theta$ of degree n and order m :

$$P_{nm}(x) = \frac{(1-x^2)^{m/2}}{2^n n!} \frac{d^{n+m}(x^2-1)^n}{dx^{n+m}} \quad (\text{A6-3.12})$$

By convention, when modelling planetary gravitational potential, equation A6-3.11 is rewritten in a slightly expanded form:

$$U(r, \theta, \phi) = \sum_{n=0}^{\infty} \left(\frac{a}{r} \right)^{n+1} J_n P_{n0}(\cos\theta) + \sum_{n=1}^{\infty} \sum_{m=1}^n \left(\frac{a}{r} \right)^{n+1} (C_{nm} \cos(m\phi) + S_{nm} \sin(m\phi)) P_{nm}(\cos\theta) \quad (\text{A6-3.13})$$

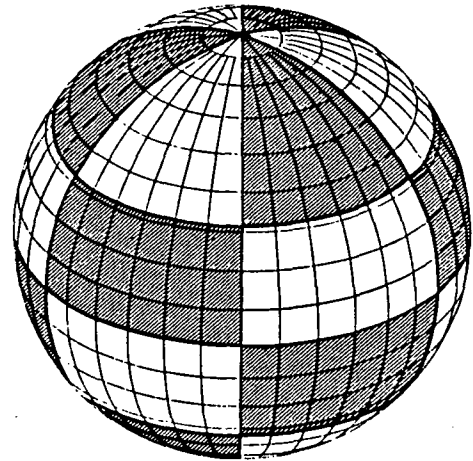
where $J_n \equiv C_{n0}$. In this form, the first term represents the "zonal harmonics", those which are independent of longitude. They are identified physically by noting that the n^{th} degree polynomial in $\cos\theta$ has n zeroes, and so changes sign n times, on the domain from 0 to π .



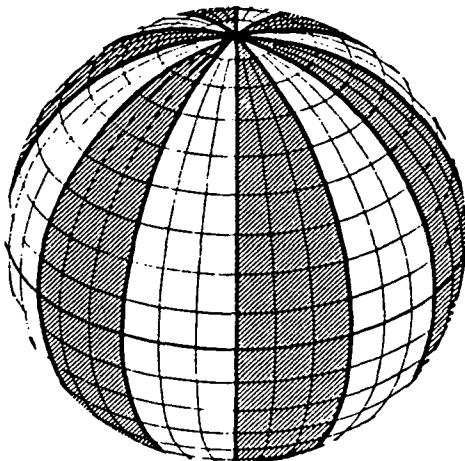
Zones for $P_6(\cos\theta)$ Spherical Harmonics

In the first of the accompanying figures taken from [Wertz, 84], alternating latitudinal bands indicate these zonal sign changes.

The second term of equation A6-3.13 represents "tesseral harmonics", which have $n-m$ zeroes on the θ domain and $2m$ zeroes on the ϕ domain (from 0 to 2π).



$P_{63}(\cos\theta)\cos 3\phi$ Showing Alternating Positive and Negative Tesseral Harmonics



$P_{66}(\cos\theta)\cos 6\phi$ Showing Tesseral Pattern Reduced to Sectoral Pattern

For $n = m$, the tesseral (tiling) pattern reduces from an alternating one to the "sectoral harmonics".

The spherical harmonic expansion can be considered as a form of Fourier series, using which an infinite sum of weighted, increasingly fine periodic functions can reproduce any bounded function. Using this model, the potential of a planet's gravity field can be numerically described in detail. Satellite geodesy is the science of measuring slight changes in the acceleration of orbiting objects, and then generating from them a catalog of the J_n , C_{nm} and S_{nm} coefficients for planetary bodies.

The advantage of writing the eigenvalue solution in the expanded form of equation A6-3.13 is that the zonal coefficients J_n represent directly important features of a planet. For instance, J_0 is simply $-GM/a$ and thus models the essential gravity feature of the planet---that is, its massive presence. The J_1 coefficient represents an offset of the center of mass from the geometrical center (where we have taken the coordinate origin). J_2 models the largest portion of an equatorial bulge, caused principally by a rotating planet's oblateness. This coefficient is the best-known because its size determines the rate of nodal regression of a satellite's orbit (Earth's J_2 is about 3 orders of magnitude larger than any other coefficient). Finally J_3 describes the largest portion of a planet's gravitational pear-shapedness; thus at Mars a major geological asymmetry between northern and southern hemispheres shows up in its J_3 coefficient, and causes large altitude variations for orbiting satellites.

Although the magnitudes of coefficients do not in general depend on their place in the infinite series, the "geometrical attenuation factor"

$$R(r) = \left(\frac{a}{r}\right)^{n+1}$$

which multiplies every term insures that higher-order terms are efficiently filtered out by increasing distance. While gravitational scientists lament the inability of even low-altitude satellites to

yield data on finer terms than $n \approx 20$, satellite engineers cheer the radius-dependent insensitivity to these higher spatial frequencies, as it allows a smoother orbital ride.

Appendix A6-4 Worst-case estimate of altitude variations at Mars
for the pentagonal resonator satellites.

The geometric attenuation factor which multiplies the J_3 zonal harmonic in the planetary gravitational potential (Appendix A6-3) is:

$$\left(\frac{a}{r}\right)^4$$

which, being larger than any subsequent analogous term, insures that this perturbation will dominate the higher-order effects. Gravitational force is proportional to ∇U ; the radial partial derivative alone will introduce another factor of r in the denominator, so force (and acceleration) effects of the J_3 term will diminish proportional to r^{-5} .

Such scaling gives us a way to adapt results from other satellites to our case. The mission for 1992 which has come to be called Mars Observer is planned for a low-altitude (361 km) near-polar circular orbit; in fact one primary mission objective involves studying the high-frequency gravity field of the planet. Various sources report predicted spacecraft altitude variations due to J_3 ranging from 40 km [Albee, 87] to 70 km [ESA, 82] to 130 km [JPL, 83]. Evidently no one will really know until the craft itself finds out. Let us choose a conservative figure of the right order, say 100 km, for these variations.

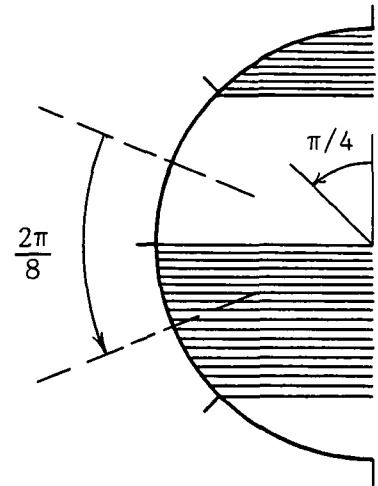
Orbital radius for Mars Observer is $361 + 3398 = 3759$ km. Since the orbital radius for a pentagonal resonator at Mars (Appendix A6-1) is 4287 km, the scaling factor due to geometrical attenuation is:

$$\left(\frac{3759}{4287}\right)^5 = .5183$$

Thus we would expect worst-case altitude variations of order:

$$(.5183) 100 \approx 52 \text{ km}$$

The third-order zonal harmonic involves three sign changes over the interval $0 < \theta < \pi$ (Appendix A6-3). The satellite time-of-flight from the center of one zone to the center of the next (1/8 the orbital period) provides an estimate of the time scale over which we could expect the maximum altitude changes to occur. For a circular orbit with semi-major axis a , the orbital period TP is:



$$\begin{aligned} \text{TP} &= \frac{2\pi}{\sqrt{\mu_\sigma}} a^{3/2} \\ &= \frac{2\pi}{(4.305(10^4))^{1/2}} (4287)^{3/2} = 8500 \text{ s} = 142 \text{ min} \end{aligned}$$

Therefore we might expect the 52 km altitude variation to occur over a time of order:

$$\frac{142}{8} \approx 18 \text{ min}$$

Appendix A6-5 Estimate of satellite altitude variation and overall system gain for square and triangular resonators in Martian sun-synchronous subsolar orbits.

Square Resonator

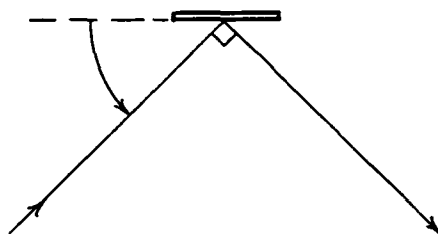
From Appendix A6-1, the orbital altitude is 4904 km. The geometric attenuation factor for J_3 force variations is then:

$$\left(\frac{3759}{4904}\right)^5 = .2646$$

Thus we estimate altitude variations on the order of:

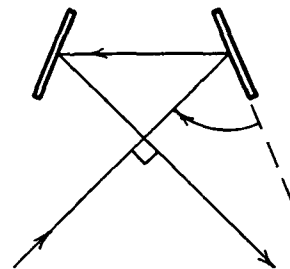
$$(.2646) 100 = 26 \text{ km}$$

To evaluate system gain to first order, we must combine the lossless minimum single-circuit gain from Figures 5-7 and 6-2 with an estimate of reflection losses around the ring. We will presume the mirrors have normal-incidence reflectivities at $10.6 \mu\text{m}$ of .995 (Chapter 7). Let us assume for now that non-normal decrements from this will follow a cosine dependence. For the square resonator we recognize two possible reflection geometries for the relay stations:



45° incidence angle
complement

4 mirrors in ring



67.5° incidence angle
complement

8 mirrors in ring

For the 4-mirror geometry, we expect:

$$\text{reflection loss} \approx \frac{4(.005)}{\cos 45^{\circ}} = 2.8 \%$$

For the 8-mirror case, we expect:

$$\text{reflection loss} \approx \frac{8(.005)}{\cos 22.5^{\circ}} = 4.3 \%$$

The 4-mirror system, being simpler and having lower reflection loss around the ring, is preferable. The minimum lossless gain for a 4-station resonator (Figure 5-7) is 100 % of the available single-pass value, which we have seen is about .058 at Mars (Figure 6-2). Thus the lossy available system single-circuit gain is:

$$1 (.058) - .028 = 3 \%$$

Ignoring other system losses, this value represents the amount of circulating laser energy we can couple out and beam into space.

Triangular Resonator

From Appendix A6-1, the orbital altitude is 6936 km, so the geometric attenuation factor becomes:

$$\left(\frac{3759}{6936} \right)^5 = .0468$$

Altitude variations will be on the order of:

$$(.0468) 100 = 5 \text{ km}$$

With a triangular geometry, incidence angles less than 60° are not possible, so the minimum loss would result from orbiting just

three mirrors:

$$\text{reflection loss} \approx \frac{3(.005)}{\cos 30^0} = 1.7 \%$$

Minimum lossless gain for a 3-station resonator (Figure 5-7) is about 85 % of the nominal .058 Martian available gain, so the lossy available system single-circuit value would be:

$$. (.85)(.058) - (.017) = 3.2 \%$$

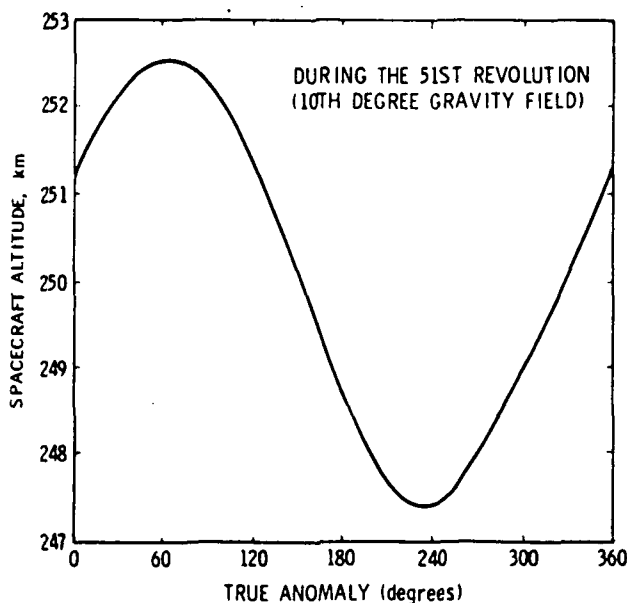
Comparing the Two

Since the difference in lossy single-circuit gain for these two resonator geometries is less than any defensible uncertainty in the gain numbers, we can consider them equal on the basis of laser performance alone. However, the fact that triangular resonator satellites would pop up and down only 1/5 as much as those in the square resonator (allowing the chance to compensate for this motion and keep their line of sight within the mesospheric inversion layer) favors the 3-station geometry at Mars.

Appendix A6-6 Estimated altitude variations of pentagonal resonator satellites at Venus.

Ballpark Method

Vijayaraghavan [84] has predicted analytically orbital perturbations for the Venus Radar Mapper (now called Magellan) due to gravitational harmonics up to 10^{th} degree and order. Primarily concerned with long-term periapsis altitude variations of its highly eccentric (0.3750359, with semimajor axis $a = 10082$ km) orbit, he takes advantage of orthogonality relations to integrate analytically his perturbation equation in full-orbit steps at periapsis --- that is, at true anomaly intervals of 2π . Maintaining that the procedure can nonetheless apply "for all values of the true anomaly", he verifies with numerical integration a 5 km "altitude" variation during the 51st orbit.



The unexplained relevance of this calculation is ambiguous since in fact during any orbit this spacecraft's altitude will vary by:

$$\frac{10082}{1 - .375} - \frac{10082}{1 + .375} \approx 8800 \text{ km}$$

due to its high eccentricity, not by the:

$$252.5 - 247.5 = 5 \text{ km}$$

indicated by the graph.

Presumably the author means that the periapsis altitude will occur variably, as shown, depending on the true anomaly location of periapsis during the 51st orbit. That exact location would of course derive

from the satellite's complete orbital history. We can therefore use his result as a measure of the altitude variation experienced by a circular orbit at that 250 km altitude, due to the 10th degree gravity field. The full 5 km change shown takes half the orbit, and so appears to model a center-of-mass offset, corresponding to $n = 1$ in the potential expansion (Appendix A6-3), rather than 10th degree roughness. Thus by analogy with Appendix A6-4, we will use a scale factor with $n + 2 = 3$ to transform Vijayaraghavan's result to our situation:

$$\left(\frac{6052 + 250}{7641} \right)^3 = .5609$$

We would expect an altitude variation of:

$$(.5609) 5 = 2.8 \text{ km}$$

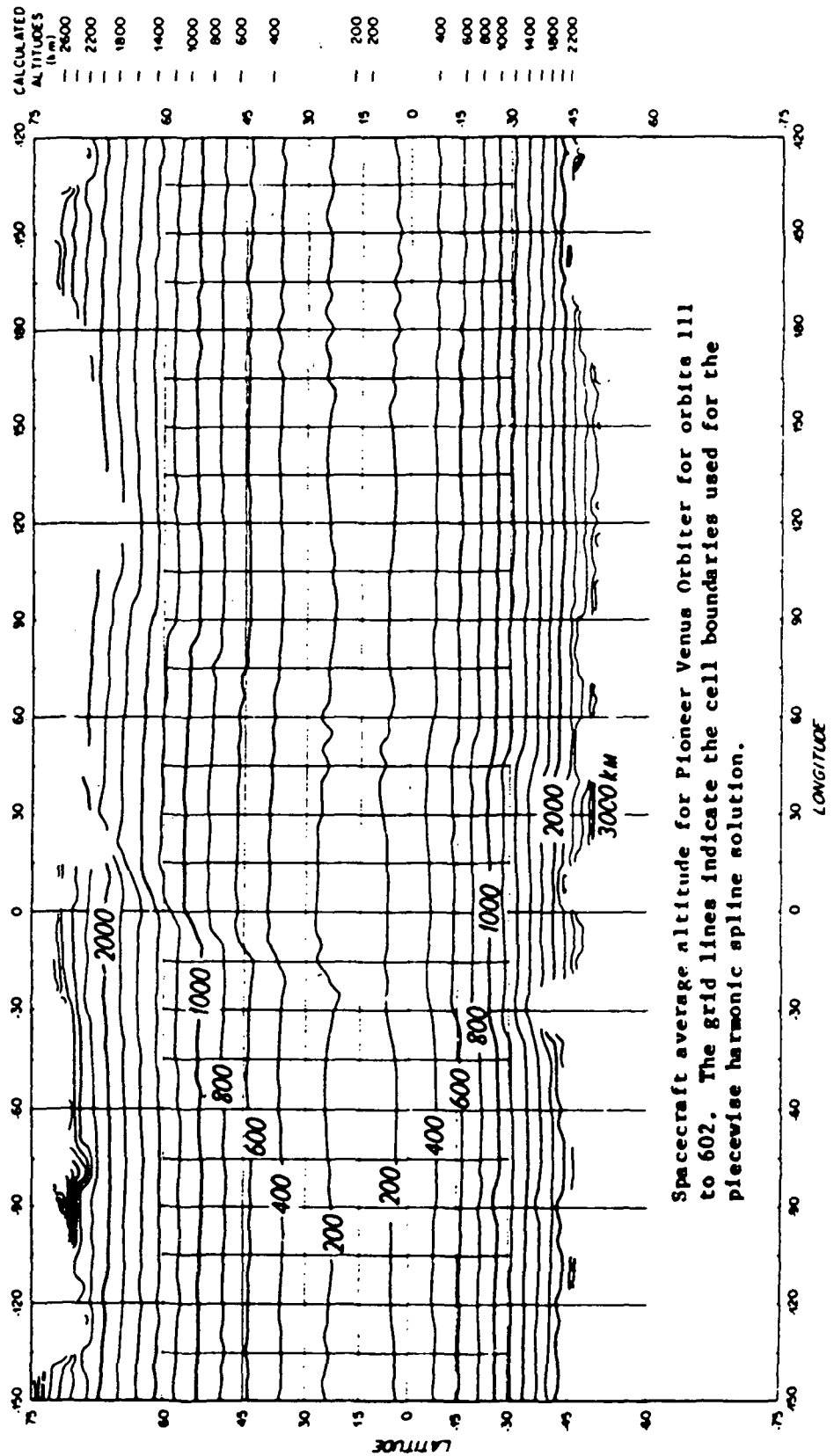
We naturally regard this value as soft, given its obscure derivation. Still, such a theoretical result is useful to corroborate the order of magnitude of a better method.

Better Method

Bowin [85] presents an extremely useful graph (reproduced on the next page) made directly from Pioneer Venus Orbiter tracking data, in which calculated altitudes for 492 separate orbits are plotted versus planetary coordinates. Because of Pioneer's long mission duration, the graph is a map, from which the effects of a highly elliptical orbit have already been removed, of gravitational bumpiness over much of Venus as experienced repeatedly by a real satellite.

We concentrate on the equatorial traces, both because our pentagonal resonator orbit is essentially equatorial and also because these traces are at low altitudes, where gravitational effects dominate solar effects

ORIGINAL PAGE IS
OF POOR QUALITY



and are measurable on the logarithmic scale. First we note the persistence but attenuation of high-frequency ripples with increasing altitude. Next we recognize that the most insistent variations (those repeated over a large range of altitudes and latitudes) occur around planetary longitudes of 60° and 180° . Near the equator, these longitudes coincide with the western edges of the two parts of Venus' largest highland, Aphrodite Terra (Figure 6-8). The fact that gravity perturbations echo clearly the continental margins, but that the continental interior itself is gravitationally as smooth as, and indistinguishable from, the lowland plains, corroborates the statement that Venus' continents are isostatically compensated. Evidently our satellites will be disturbed by the planet only when crossing over Aphrodite's edges.

The worst of the altitude variations is the one around $\phi = 60^\circ$, which takes about 15° of longitude (true anomaly for a circular equatorial orbit) from peak to peak, and consists at the 200km nominal altitude of a roughly 30 km excursion. Now 15° is $\pi/12$, so this high-frequency anomaly cannot be modeled by a term of degree lower than 12 (see illustrations to Appendix A6-3). Thus our scale factor must be of degree $n + 2 = 14$:

$$\left(\frac{6052 + 200}{7641} \right)^{14} = .06$$

Consequently we expect altitude variations of order:

$$(.06) 30 = 1.8 \approx 2 \text{ km}$$

for the resonator satellites, a value in good agreement with our earlier soft value.

Appendix A6-7 Astrodynamical reference values for a pentagonal
planetary resonator at Venus.

Starting values are taken from [Wertz, 84], [CSC, 77], [Hunten
et al, 83], [Deming & Mumma, 83].

The Venus Orbit

Orbital Eccentricity	e	.006787
Semimajor Axis	a_v	108.2(10 ⁶) km
Perihelion Distance	$r_p = a_v / (1 + e)$	107.5(10 ⁶) km
Aphelion Distance	$r_a = a_v / (1 - e)$	108.9(10 ⁶) km
L1 Distance from Venus	(at $r = a$)	1.007993(10 ⁶) km
L2 Distance from Venus	(at $r = a$)	1.014292(10 ⁶) km
Sidereal Orbital Period		224.7 d
Mean Orbital Speed		35.03 $\frac{\text{km}}{\text{s}}$
Mean Solar Flux		2.60 $\frac{\text{kW}}{\text{m}^2}$

The Planet Venus

Equatorial Obliquity		3°
Sidereal Rotation Period		- 244.3 d
Equatorial Radius	R	6052 km
Ellipticity	$(R_e - R_p)/R_e$	0.0
Gravitational Constant	μ_{ϕ}	$3.248588(10^5) \frac{\text{km}^3}{\text{s}^2}$
Effective Temperature	T_e	231 K
Altitude of Mesospheric Inversion		128 - 138 km
Altitude h of Peak Integrated Tangential Emission		130 km
Kinetic Temperature T at Inversion Altitude		195 K

The Pentagonal Resonator

Number of Relay Stations		5
Orbital Radius	$r = \frac{R + h}{\cos 36^\circ}$	7641 km
Orbital Altitude	$h_s = r - R$	1589 km
Orbital Speed	$v_{cs} = \left(\frac{\mu_\oplus}{r} \right)^{\frac{1}{2}}$	$6.52 \frac{\text{km}}{\text{s}}$
Orbital Period	$TP = \frac{2\pi}{\sqrt{\mu_\oplus}} r^{3/2}$	7363 s
		= 122.7 min
		= 2.045 h
Satellite Separation	$l = 2(R + h)\tan 36^\circ$	8983 km
Light Propagation Time Around the Ring		
	$\tau \approx \frac{5(8983)(10^3)}{3(10^8)}$	0.150 s
Ring Displacement During τ		
angular	$\omega \approx \Omega\tau = \frac{2\pi}{7363} (.150)$	$1.278(10^{-4}) \text{ rad}$
tangential	$s = r\omega$	976 m

References

- Brij N Agrawal- Design of Geosynchronous Spacecraft
(Prentice-Hall, 1986).
- Arden L Albee- "US and Soviet Plans for Exploration of Mars -
Return to the Red Planet" Engineering Colloquium at NASA
Goddard SFC (4 May 1987).
- M P Ananda, W L Sjogren, R J Phillips, R N Wimberly, B G Bills-
"A Low-Order Global Gravity Field of Venus and Dynamical
Implications" Journal of Geophysical Research Vol 85
No A13 pp 8303-18 (30 Dec 1980).
- Graeme Aston- (b) [op cit Chapter 9].
- G Balmino, B Moynot, N Vales- "Gravity Field Model of Mars in
Spherical Harmonics up to Degree and Order Eighteen"
Journal of Geophysical Research Vol 87 pp 9735-46
(1982).
- Roger R Bate, Donald D Mueller, Jerry E White- Fundamentals of
Astrodynamics (Dover, 1971).
- Jürgen Blunck- Mars and Its Satellites 2nd ed
(Exposition, 1982).
- Carl Bowin- Gravity Anomaly Map of Mars and Moon and Analysis
of Venus Gravity Field: New Analysis Procedures
Final Report NASA-CR-174295 (1985).
- B Fong Chao & David P Rubincam- On Seasonal Variations of
Mars' Gravitational Field NASA TM 87807 (1987).

Computer Sciences Corporation- Tables of Astronomical Constants and Libration-Point Parameters for Selected Solar System Objects NASA GSFC Contract NAS 5-11999
Task Assignment 570 CSC/TM-77/6025 (Jan 1977).

E Cutting, J H Kwok, S N Mohan- "Venus Radar Mapper (VRM) Mission" AIAA 84-0212 (12 Jan 1984).

Drake Deming & Michael J Mumma- [op cit Chapter 2].

P B Esposito, S Demcak, M L Santee- "Mars Observer Orbital Accuracy Analysis" AIAA Paper 86-2057 (20 Aug 1986).

P B Esposito, W L Sjogren, N A Mottinger, B G Bills, E Abbott- "Venus Gravity: Analysis of Beta Regio" Icarus 51 pp 448-59 (1982).

European Space Agency- Kepler: An Interdisciplinary Mars Orbiter Mission, Report on the Phase A Study ESA SCI(82)5 (Dec 1982).

B F Gordiyets & V Ya Panchenko- [op cit Chapter 2].

William K Hartmann- [op cit Chapter 2].

D M Hunten, L Colin, T M Donahue, V I Moroz (eds)- Venus (University of Arizona, 1983).

Jet Propulsion Laboratory- Mars Geoscience Climatology Orbiter 1990 JPL Technical Definition Review (26 May 1983).

Marshall Kaplan- [op cit Chapter 8].

- J H Kwok- "Long-Term Orbit Prediction for the Venus Radar Mapper Mission Using an Averaging Method" AIAA 84-1985 (22 Aug 1984).
- P L Lawing, D A Dress, R A Kilgore- Potential Benefits of Magnetic Suspension and Balance Systems NASA TM-89079 (Feb 1987).
- H Masursky, E Eliason, P G Ford, G E McGill, G H Pettengill, G G Schaber, B Schubert- "Pioneer Venus Radar Results: Geology from Images and Altimetry" Journal of Geophysical Research Vol 85 No A13 pp 8232-60 (30 Dec 1980).
- S N Mohan & P B Esposito- "Venus Radar Mapper Orbit Accuracy Analysis" AIAA 84-1986 (22 Aug, 1984).
- G H Pettengill, E Eliason, P G Ford, G B Lorient, H Masursky, G E McGill- "Pioneer Venus Radar Results: Altimetry and Surface Properties" Journal of Geophysical Research Vol 85 No A13 (30 Dec 1980).
- William L Sjogren- "Mars Gravity: High-Resolution Results from Viking Orbiter 2" Science Vol 203 (9 Mar 1979).
- W L Sjogren, R J Phillips, P W Birkeland, R N Wimberly- "Gravity Anomalies on Venus" Journal of Geophysical Research Vol 85 No A13 pp 8295-302 (30 Dec 1980).
- Sean Solomon- "Tectonic Evolution of Venus and the Earth" Scientific Colloquium NASA Goddard SFC (29 May 1987).
- Gordon E Taylor- "Oblateness of the Atmosphere of Mars" Nature Vol 264 (11 Nov 1976).

Chauncey Uphoff- "Future Planetary Orbit Missions" in:
Natural and Artificial Satellite Motion (Paul E Nacozy &
Sylvio Ferraz-Mello, eds) (University of Texas, 1979).

Petr Vanicek & Edward J Krakiwsky- Geodesy: the Concepts
(North-Holland, 1982).

A Vijayaraghavan- "An Analytic Solution for the Orbital
Perturbations of the Venus Radar Mapper Due to
Gravitational Harmonics" AIAA-84-1995 (22 Aug 1984).

James R Wertz (ed)- Spacecraft Attitude Determination and
Control (Reidel, 1984).

A F Whitaker, S A Little, R J Harwell, D B Griner, R F DeHaye,
A T Fromhold Jr- "Orbital Atomic Oxygen Effects on Thermal
Control and Optical Materials: STS-8 Results" NASA Tech
Briefs MFS-28084 (Apr 1987).

B G Williams & N A Mottinger- "Venus Gravity Field: Pioneer
Venus Orbiter Navigation Results" Icarus 56
pp 578-89 (1983).

CHAPTER 7

THE OPTICAL PATH

Chapter Abstract - Single-surface vertex stations confine a 1 km diameter laser beam. Cooperative, individually precise mirrors control the cavity modes without NOPC; required actuator resolution is 62 nm. A diffraction grating selects the P(12) line for oscillation, polarizes the field and couples out 180 kW. System gain exceeds the distributed loss budget. The Station 1 constellation conditions a 10 m intermediate beam, directing it out of the resonator to Venus' collinear libration points; required pointing accuracy is 5 nrad. L1 and L2 Stations can accommodate a variety of modulator types to impress the signal, apply the proper divergence and aim at stellar targets.

The orbiting equatorial pentagon, together with "fixed" diametric libration points, establishes a gross armature within which continuous laser transmission could occur from Venus. Next we begin filling in details of this framework to build a plausible arrangement of hardware and control which can comprise that laser system. This chapter combines communication requirements from Chapters 1 and 3 with laser principles from Chapter 2, given the geometry developed in Chapters 5 and 6, to derive a workable optical scheme for the Venusian planetary laser. We navigate a course through many physical constraints, transforming them into engineering specifications which must and can be met by the subsystems outlined in subsequent chapters.

Cavity Diameter

The fundamental design detail is a proper cross-sectional size of the intracavity laser beam. Equation A3-3.7 shows clearly that achievable data transfer rate varies directly with the fourth power of cavity diameter; we therefore want as large a diameter as possible. Countering this is the certainty that whatever optical, dynamic, control, fabrication and reliability problems the system incurs will increase by at least the square of cavity diameter (proportional to area). More quantitatively, in Chapter 6 we saw that diameters larger than about 5 km cannot guarantee continuous operation in any case. Later in this chapter, however, we find that plane reflector diameters smaller than 100 m introduce unacceptable intracavity diffraction losses. The somewhat arbitrary choice of a 1 km diameter thus seems appropriate for baseline design; within the physically admissible size range, it is

large enough to exploit the available kilometer-scale inversion layer without being profligate.

Increasing cavity diameter to the 5 km maximum would allow a first order 625-fold performance improvement over our baseline. We should recognize also that even a 1 km laser beam extracts solar-pumped energy from only one five-millionth the volume of the mesospheric gain shell. Thus merely by engineering a planetary-scale system, we are not in any noticeable way engineering the planet itself; a 1 km laser is not really large at all, compared to the untapped remainder of its energy source.

Satellites with linear dimensions of order 1 km are not inherently unreasonable by current planning standards. Large space structures (LSS) even exceeding this size have been seriously proposed for many years [Bock, 79] [O'Neill, 78]. What is unusual about our satellites is that, being cavity reflectors for a laser, they must operate with optical accuracy despite their size and separation, a mission quite beyond those yet planned. All optically accurate space devices officially envisioned by NASA have diameters smaller than 100 m [Soosaar, 84], and tenable SDI optical diameters appear to be only a tenth as large [Smith, 87]. This mismatch then comprises our most general engineering specification: effective use of what Venus provides requires kilometer-scale satellites of telescope quality.

Basic Vertex Stations

Since each resonator vertex station always has a clear line of sight to one of the distant diametric libration points L1 and L2, only one of the five stations need couple laser energy out of the cavity. Specializing one for that purpose allows the other four to be simpler and identical. By characterizing these latter vertex stations, labeled 2, 3, 4, and 5, we bound the basic laser cavity.

As mentioned already in Appendix A6-5 and shown in Figure 7-1, essentially two ways of folding the beam at a vertex are possible. Intuition suspects that the extreme rays tracing closed paths around a ring bounded by angled mirrors of non-zero size would differ in length (since, after all, circumferences corresponding to different radii are different). That would mean, however, that each infinitesimally adjacent transverse portion of the cavity would support a different resonant wavelength, precluding spatial coherence of the beam; in fact, this does not happen. A plane wave reflecting off a plane mirror will remain plane, so that spatial coherence is independent of the number of reflections required by the cavity. Appendix A7-1 offers a rigorous proof for the skeptical. Since geometrical optics alone cannot choose between the two vertex geometries, we must examine their relative merits in some detail.

At first the double-surface method seems advantageous. Enclosing a short portion of the laser beam in a vertex "pocket" would facilitate rapid, accurate cavity length adjustment, necessary for emission-linewidth narrowing as we shall see later. Also, near-normal incidence angles (180° for the pentagon) maximize specular reflectivity for general polarizations. Neither of these features is a robust benefit, however.

All stations must actively control cavity length together simultaneously since any one is by itself capable of spoiling the laser; a "tuning pocket" would be superfluous. As for the advantage of near-normal incidence, general relations between incidence angle and specular reflectivity are not readily available since empirical details vary widely. Reflectivity at $10.6\text{ }\mu\text{m}$ of gold or silver with a thin-film reflection-enhancing interference filter is typically at least 0.995 for normal incidence [Two-Six]. Following our conservative approach from Appendix A6-5 of increasing this 0.005 reflection loss by the cosine of incidence angle, we might expect for the double-surface configuration a 4.2 % total loss from the 8-surface sequence of Stations 2, 3, 4 and 5 on each pass.

It turns out that reflectivity for general polarizations is irrelevant to our laser though, since our wavelength selector will enforce linear polarization anyway. If we arrange that selector properly we can count on higher reflectivity. The industry value for enhanced gold or silver with s-polarized $10.6\text{ }\mu\text{m}$ light increases to a minimum of 0.997 at 45° incidence angle [Two-Six]. Now the single-surface station type requires incidence angles of 54° , which should improve further upon this higher value, as it represents an even more "glancing" ray. Additionally, only four surfaces are required for the four basic stations with this method, so even using a conservative 0.997 reflectivity, we would expect no more than 1.2 % total reflection loss per pass through the sequence of Stations 2, 3, 4 and 5. Since we have only 10 % gain per circuit to begin with, the single-surface configuration represents a substantial performance benefit.

Other important advantages (besides the obvious one that it reduces the number of expensive satellites and consequent sources of possible system failure) favor the single-surface method for basic vertex stations. Presuming reflexive mass symmetry, single-surface satellites would fly "principal axis planet-oriented", yielding no cross products in their inertia tensors (Chapter 8). Thus they would be conditionally stable with respect to gravity gradient torques, "of primary importance for large, massive platforms" in close orbits [Woodcock, 86]. Geometrically non-principal axis orientations (like the two-surface satellites, to first order) would require structure, mass asymmetry, or severe logistics penalties to compensate secular gravity gradient torques.

Additionally, the optical surfaces of the stations must of course be completely exposed, with unobstructed lines of sight to the two adjacent stations. Keeping those precious surfaces parallel to the velocity vector minimizes the ram flux of reflectivity-degrading contaminants and abrading particles (because the projected optical area normal to the velocity is zero). Additionally, keeping them facing down toward the planet minimizes the space flux of damaging micrometeors that they will see. Locations near a planetary body incur a meteor impact penalty because the planet's gravity has a focusing effect on passing space debris. Appendix A7-2 shows that we might expect about 1.3 times the deep space flux at our orbital altitude. However, planetary bodies also block a portion of the 4π sr field seen by objects in close orbits, so Appendix A7-2 goes on to predict Venus shielding our satellites from about 19 % of the inflated natural flux just estimated, bringing it back down to 1.05 the deep space value overall. Significantly though, the necessary asymmetry of single-surface vertex stations lets their vulnerable optical surfaces face the planetary shield directly, leaving their backup structure and control hardware to absorb most of the

gravitationally magnified space flux. By thus exploiting directly the anisotropic particle flux distribution near Venus, the single-surface configuration protects the optical surfaces most effectively.

The 54° incidence angle means, of course, that since the reflector surface projected normal to the beam must still be 1 km across, the reflector itself must be an ellipse with minor axis roughly 1 km and major axis roughly 1.7 km. This increased area penalty must necessarily be less, though, than the area and complexity penalties introduced by the double-satellite alternative. Our second major engineering specification, then, is for single-surface elliptical plane reflectors at Vertex Stations 2, 3, 4 and 5.

Phase Conjugation

What can be done to provide optically accurate plane reflectors over a kilometer across? Diffraction-limited performance requires mirror surfaces accurate to $\lambda/20$, or about $0.5 \mu\text{m}$ in our case; certainly such tolerance over a scale of kilometers -- one part in $2(10^9)$ -- will not be attained easily. Before exploring strictly engineered solutions, we recall from Chapter 2 that various techniques of non-linear phase conjugation, when applied to laser resonators, can obviate the need for perfect reflectors altogether. Naturally we would prefer the elegant physics of a PCR to the baroque intricacy of a Rube Goldberg machine.

Stimulated Brillouin Scattering would permit the most straightforward geometry, requiring only a proper medium to act as the PCM. Unfortunately, such media would typically be gases

at pressures of many atmospheres; a kilometer-scale vessel to contain those pressures and be transparent to infrared radiation would be hard to invent. Focusing the beam down would reduce the required PCM size (and help meet the SBS threshold) if the resulting power density did not cause medium dissociation or IR window damage. But we know that SBS Stokes-downshifts its scattered light, making the conjugated beam useless to a narrow-line resonator after only one reflection. Furthermore, simply retrodirecting the incoming beam is unacceptable for our co-orbiting case; even if the penultimate mirror sent the beam to where the PCM would be by the time the light got there, the PCM would send the conjugated beam back to the source mirror's original location, missing its updated position by $2/3$ km!

The greater versatility of Four Wave Mixing makes it the more commonly used technique for experimental PCRs. Most mixing is done in photorefractive crystals, but again, high power densities would require quite large crystals for our use. We might with fairness imagine these manufactured with advanced microgravity techniques, but the transmission losses of such materials would undoubtedly soak up too much of our oscillator gain. However, inverted gaseous CO_2 (which Venus obviously provides) has been used for FWM. We know that some unique three-dimensional pump-detuning geometry exists to effect any modest combination of frequency and angular offset, so it might seem possible to develop a geometry to perform FWM for the planetary laser. Several tough problems immediately arise, only a few of which we will examine.

First of all, the published data on efficient FWM in inverted CO_2 are scant, so basing our laser's operation on them seems reckless if not absolutely necessary. The only available interaction regions are those tangent mesospheric volumes already being depleted by the laser itself at the solar pumping

rate. The laser would therefore have to "share" its inverted medium with the FWM process; by itself this is not bad (maximum FWM efficiency occurs for pumps and probes of equal intensity), but it would halve the available gain, extinguishing the laser. Furthermore, it is questionable whether 50 % efficiency could be attained with a virtually collinear pump-probe angle, because with this geometry one of the induced gratings tends to become washed out.

Although the extra satellites to define a FWM geometry could probably be formation-flown using small perturbative forces, at least three sets would be required for continuous conjugation since up to three interaction regions suffer darkside passage (and no inversion!) simultaneously. Finally, plane (or at least phase-conjugate) pump beams are necessary, efficient FWM requires laser operation in single longitudinal and transverse modes, and all the aberrated light must enter the interaction region to be conserved; these prerequisites call for a degree of mirror control that FWM was supposed to preclude in the first place.

While non-linear phase conjugation unquestionably enables many marvels, planetary lasers do not yet seem to be among them. An enormous gap separates fascinating laboratory results from particularly large-scale applications. The remainder of this project will therefore seek more familiar optical control solutions. NOPC technology is still quite new, though; by corollary to Clarke's First Law [Clarke, 84], we cannot pretend that advancing knowledge could never skirt all the dilemmas raised in this section. A rational view would be that, with the feasibility of planetary lasers based on other means, NOPC might still someday make such devices easier, perhaps cheaper, perhaps better.

Mirrors

Chapter 2 surveyed some elaborate mechanical techniques, two of which worked well in limited sizes, for attaining diffraction-limited mirror performance at $10.6\text{ }\mu\text{m}$ based on controlled surface deformation. The device used by Stephens and Lind [78] achieved good results with a thin metal membrane figured by actuators at 2 cm spacing, a value determined by the wavelength, intrinsic membrane stiffness, and actuator strength. Such an arrangement applied to our case would mean on the order of $4(10^9)$ actuators for each satellite, baroque indeed.

The other prime candidate was Fisher's self-referenced IPL system [85], which shares some strengths and weaknesses with NOPC techniques. Although robust because monolithic, and not inherently size-limited, its self-referenced operation requires projecting a phase error intensity pattern onto its rear face. Tapping the necessary information from the incoming light with a full-field beam splitter contributes an intracavity loss which we simply cannot afford; in any case, keeping both the front and back of a 1 km mirror unobstructed seems unlikely. Ultimately the system could not be strictly self-referenced anyway, because each satellite in the planetary resonator has to control its surface cooperatively with its fellows. The device thus becomes essentially a high-resolution deformable mirror, and its $10\text{ }\mu\text{m}$ channel spacing would mean of order 10^{16} individual elements per satellite for us. This is not so much baroque as surreal.

The opposite extreme would be to posit monolithic mirrors so massive that their stiffness prevented unacceptable deviations of the $\lambda/20$ surface miraculously machined across their square kilometers of area. One might argue that the resource commitment necessary to build a planetary laser in the

first place must include access to sufficient materials to accomplish this. Indeed, were the micromanufacturing skill available, such an approach would seem to merit serious attention, at least for Vertex Stations 2, 3, 4 and 5. Before diving into material and structural properties, however, we note (as preview) that other optical elements of the planetary laser system, and its overall performance, can benefit from adaptiveness that, once fabricated, a megamirror could not provide. Further, environmental perturbations will require continual, if subtle, cavity length and reflector pointing readjustment. μm linear and μrad angular fine-tuning of an object a kilometer across and massing perhaps tens of millions of metric tons is well beyond the accuracy of any system we know with the control authority to move it. We will not consider megamirrors further for the large cavity components.

Having dismissed as unfeasible the preceding extremes, we naturally try an intermediate approach. A contemporary method for achieving the performance of optical telescope mirrors larger than can be made monolithically uses segmented mirrors. Both the new 10 m Keck IR telescope at Mauna Kea and the planned 30 m space-based IR Large Deployable Reflector are examples. Instead of trying to make the entire surface excellent at once, many excellent smaller mirrors are made to work together. Individual diffraction-limited mirrors, when controlled cooperatively to act as though they were one large mirror, yield diffraction-limited performance for their total aperture. Called Huygens' principle, this clever application of physical optics lets telescope performance be limited by control technology rather than by mirror size. Picturing our large satellite reflectors as cooperative arrays of more conventionally sized mirrors, looking something like planar versions of an insect's compound eyes, brings them into the realm of calculable feasibility.

Chapter 8 will explore ways of moving and controlling such mirrors, but here we investigate the mirrors themselves, based on [Yoder, 86]. The highest performance optics are "first-surface" mirrors, having their reflecting surface in front of a structural substrate so that no light is lost by transmission. Metallic coatings are typically evaporated onto a ground and polished glass substrate (still the most common). Electroless plated nickel currently sets the standard for durable, high-quality reflection; it is easily machined to extreme smoothness with Precision Diamond Turning (PDT) techniques, which maintain $0.05\text{ }\mu\text{m}$ dynamic tolerances between the workpiece and a single point diamond tool.

Mirror surfaces are prey to three imperfection regimes. Figure errors represent the overall deviation of a reflecting surface from its mathematically optimal shape, due to residual fabrication errors or environmental influences; as mentioned earlier, figure error amplitudes greater than about $\lambda/20$ produce wavefront aberrations which limit performance. Midfrequency errors are deviations, resulting from manufacturing peculiarities, with spatial frequencies between about 10 and 250 cycles/m. Finally, surface microroughness of up to 50 cycles/mm, due to inherent material limitations, degrades specular reflectivity by scattering some of the light.

Glass is the favored ground-based mirror substrate for many reasons, not the least of which is the inertia of tradition. Relatively cheap and easy to cast, its residual stresses after polishing can be completely removed by proper annealing, contributing greatly to dimensional stability during use. The amorphous nanostructure of vitreous materials makes it possible to polish them to the lowest achievable microroughness, about 0.5 nm rms . Tailored formulations like Cer-Vit and Zerodur exhibit near zero coefficients of thermal expansion around 300 K (room temperature), minimizing the

distorting effect of temperature changes --- probably the greatest environmental variable for telescopes.

A large thermal inertia (high specific heat) also helps resist transient temperature variations, though, and a high thermal conductivity ensures rapid and even cancellation of distorting thermal gradients within the substrate. So-called athermal design emphasizes these attributes to achieve stable performance, particularly in high thermal-flux environments. Structural rigidity is also especially important for space applications, where low mass and high fundamental vibration frequencies are always sought. The materials which excel in these last three areas are not glasses but rather metals, so an entire technology has arisen developing large optical mirrors from aluminum, copper, molybdenum, and beryllium. At the cryogenic temperatures commonly used for IR detector telescopes, even metals exhibit near-zero CTE; furthermore, most metals are intrinsically highly reflective at infrared wavelengths, so machinable amorphous surface layers of base metals can be polished directly. Mirrors of Mo and Be, which do not machine well, can be coated by electroless nickel or gold to accept PDT finishes. Surface microroughness values comparable to vitreous mirrors are then achievable.

All large modern mirrors take structural advantage of careful material placement to minimize the amount of material not actually working optically. The effectiveness of this lightweighting process is judged by the percentage mass reduction from a solid slab required to meet the same optical criteria. Glass mirrors can be fabricated by building up an open-cell core to which face sheets are bonded (as was done for the 200" Hale), or cast directly over hexagonal plugs, forming honeycomb ribs, as Roger Angel does now [Thomsen, 87]. Metal mirrors, even beryllium, are generally amenable to mechanical or chemical milling [Yoder, 86], allowing optimized

lightweighting factors. "Eggcrating" the inside of both faces and then brazing them together can immediately double the total lightweighting factor for metal mirrors [NASA TB, 83].

Beryllium mirrors have been developed extensively. Although highly toxic in dust form and difficult to machine, optical grade I-70A Be possesses other quite special and favorable characteristics (it is the fourth simplest element in nature). Among the mirror metals, its average thermal expansion coefficient is intermediate. But its thermal conductivity (220 W/m-K) is second only to copper, and its specific heat (1820 J/kg-K) is the highest. Perhaps most telling, its specific stiffness (defined as Young's modulus of elasticity divided by density) is $16.4(10^7)$ Nm/kg, over 1/3 larger than silicon carbide, over 5 times larger than molybdenum (the closest mirror metal), and over 4 times larger than Cer-Vit (the most competitive glass). Thus despite its low density ($1.85(10^3)$ kg/m³) it can deliver superior structural performance, albeit at 1/3 to 1/10 the strength (microyield stress, the stress in uniaxial tension causing a permanent strain of 10^{-6} [Barnes, 77]) of various aluminum alloys.

Beryllium has a hexagonal crystal lattice rather than the cubic lattice characteristic of the other mirror metals. Consequently its thermal expansion is anisotropic. (The expansion also varies with temperature, although not as drastically as does aluminum's.) Reference values can be taken as $\alpha_a = 8.0(10^{-6})/K$ and $\alpha_c = 5.2(10^{-6})/K$, both at 200 K. Respective values at 400 K are $15.0(10^{-6})/K$ and $11.1(10^{-6})/K$ [Am Inst Phys, 72]. Typically the metal is made macro-isotropic by powder metallurgy processing [Yoder, 86]. Once ground up, the anisotropic micrograins orient randomly; the material is compacted, sintered and finally subjected to Hot Isostatic Pressing (HIP). What results is a 99.8 %

dense, isotropic and slightly stronger (probably due to strain hardening) beryllium blank which can be milled and coated. Heat treatment after each machining step is required to relieve internal stresses, as with all metal mirrors. Although the reflectivity of polished bare beryllium at wavelengths greater than about $5 \mu\text{m}$ is $> 97 \%$, scattering due to the material's intrinsic microroughness limits it there. Higher reflectivities require coating, for instance with electroless nickel or gold. The penalties are increased weight, decreased radiation tolerance, and bimetallic behavior with changing temperature.

Perkin-Elmer, the most experienced manufacturer of large IR optics, has recently developed a proprietary HIP process for making I-70A beryllium mirror blanks which yields superb results [Paquin et al, 84]. Their conservative demonstrator mirrors were 0.24 m diameter; a monolithic hexagonal honeycomb core of 12.5 % density was formed and sandwiched between facesheets 0.25 cm thick in a single operation. The 0.98 kg blanks' surface figure errors were less than $\lambda/25$ in the visible, an order of magnitude better than necessary for IR wavelengths. Perkin-Elmer concluded that "the process is now ready for application to production optics".

The researchers developed theoretical mechanical data for scaled-up versions of 17" and 36" diameters, based on an areal density of 21.1 kg/m^2 (which they claim could be halved with little performance decrement). They predict diffraction-limited surface figure accuracy for visible wavelengths, and say such mirrors need support at only three points on the back for polishing, metrology and use. That is, these mirrors can be treated as optically perfect rigid bodies; once finished and installed, only the mounting points would need control. Appendix A7-3 backs off slightly from the Perkin-Elmer optimism, applying conservative assumptions to extend their

projected size modestly. We thereby develop a basic mirror configuration for our planetary resonator stations.

The blanks are hexagonal flat sandwich slabs, with facesheets enclosing a structurally optimized honeycomb core, formed and HIP-bonded from I-70A beryllium powder. The large dimension is 3 m, the total thickness is 0.25 m, and the mass is 180 kg (including a sputtered amorphous beryllium surface for microsmoothness and a coating for high reflectivity); the fundamental frequency is about 600 Hz. Three attachment points on the back are lined up with vertices and spaced equally around a circle of 1 m radius.

Silver has a slightly higher intrinsic reflectivity at 10.6 μm than gold [Amer Phys Inst, 72], and is certainly cheaper. We have seen, however, that when these metals are coated with enhancement interference filters their reflectivities are comparable. Gold is essentially inert, whereas silver readily oxidizes when exposed to atomic oxygen (a major component of terrestrial planet thermospheres), particularly at elevated temperatures. Dielectric coatings (like a reflection-enhancing coating) can protect silver substantially, but this has only been tested for short durations, in benign thermal environments, and to moderate measurement accuracies [Whitaker et al, 87]. Even at the relatively high altitude of the Venusian resonator satellite orbit, a thin silver oxide layer can be expected to form eventually on mirror surfaces, reducing reflectivity below unacceptable levels after years of operation. Non-contacting methods of oxygen removal based on electron-beam stimulated desorption have been demonstrated [Outlaw et al, 87], so we might envision mirror-cleaning robots. For simplicity, though, we will avoid this problem entirely by specifying enhanced gold coatings instead.

Frequency Selection and Output Coupling

At this point we have actual mirrors capable of being incorporated into vertex stations to define the beam cavity. Next we must ensure that the laser oscillates on only one of the possible CO₂ rotational lines (to prevent line competition from keeping the laser below the oscillation threshold), and in only one direction around the continuous ring (to avoid frequency splitting due to ring rotation, as explained in Chapter 6). Then we have to couple out a small amount of the circulating power from the cavity, to modulate and use. We will achieve all these goals simultaneously using a diffraction grating, but first we should evaluate the alternatives.

Frequency selection involves spoiling the gain for undesired frequencies which would otherwise oscillate, bleeding energy away from the desired frequency. These can be other rotational transitions, in the case of a molecular laser like CO₂, each of which has its own gain profile separated by about 50 GHz from its neighbors, or separate cavity modes within a single 50 MHz-wide gain profile. That is, any frequency whose corresponding half-wavelength divides integrally the cavity length L , can oscillate in that cavity. If in addition that frequency lies within a laser line (gain profile), then amplification can occur upon oscillation. For a cavity of such enormous L as ours, a vast number of different frequencies might oscillate with gain, broadening the emission by dissipating the available power instead of concentrating it on one frequency. Narrowband interference filters cannot provide the required resolution even to keep unwanted lines from oscillating. By far the simplest line-selection process is to disperse the light spectrally, rejecting from the cavity the emissions of undesired lines so they cannot oscillate. Conventionally, multiline lasers use diffraction gratings for

this cavity tuning. We will return to the second problem of fine-tuning the cavity oscillation later.

Several techniques exist for getting a beam out of a laser cavity. Decreasing slightly the reflectivity of one resonator mirror to permit some transmission is the most common, but obviously can work only for mirrors on transmissive substrates and so is unavailable to us. An analogous technique common in FIR lasers is to leave a hole in one mirror of area commensurate with the desired coupling loss. For our use this technique would cause problems. First, it would establish a spatially heterogeneous intensity pattern across the intracavity beam, an effect of unknown ramifications. Also, it would necessitate leaving an unobstructed beam path behind the coupler mirror, free of structure and mechanisms. For simplicity we would prefer keeping the laser on one side of the mirror plane, and the spacecraft housekeeping functions on the other, preventing mutual interference.

An intracavity beamsplitter could remove laser energy. Conventional splitters are not reasonable for us because they include a transmissive path; again, an unobstructed path behind the splitter would be necessary. Sometimes IR lasers use a reflecting grid as a splitter --- most of the circulating power passes through the openings to the resonator mirror, while the rest is reflected by the grid surface out of the cavity. While this scheme might work for us, the output beam would of course be spatially heterogeneous. Without detailed analysis, the far-field implications of a gridded beam are unclear, but unwanted if avoidable.

Light which strikes an edge is diffracted, meaning scattered or redirected. Diffracted monochromatic light will self-interfere, producing bright and dark fringes as the waves add and cancel (Figure 7-2). The fringe spacing produced by a

slit depends on both the slit width and the light wavelength; other colors of light, scattering at different angles, yield other fringe spacings. A diffraction grating can be considered simply as an infinite set of parallel slits [Bausch & Lomb, 70] which disperses light spectrally as does a prism. Instead of exploiting the retardation of different wavelengths in a transmissive medium of non-unity refraction index, though, the grating is a wholly reflective optic based on the interference of electromagnetic waves. Figure 7-3 defines the standard diffraction grating notation, and Appendix A7-4 derives the grating equation.

A grating used such that the diffracted angle β is identical to the incidence angle α is said to be Littrow autocollimated. For a "blazed" grating ($\alpha = \beta = \theta$, the blaze angle) which is autocollimated, diffraction into the $m = 1$ order (the diffracted ray retracing the incident ray) far exceeds diffraction into other orders, because the light is reflected normal to the groove face [Bausch & Lomb, 70]. Appendix A7-4 also shows how the grating equation simplifies in this case. The Littrow mode is used extensively to tune laser cavities, since almost all of the light with proper frequency goes right back into the cavity to maintain oscillation; the rest is scattered into just a few other orders and therefore is coupled out at predictable angles, while light at all other, undesired frequencies is scattered at other angles and hence rejected. In fact, it can be proved (Appendix A7-5) that for a Littrow grating with groove spacing d between $\lambda/2$ and $3\lambda/2$, only the orders 0 and 1 are allowed. This permits the lowest possible losses; the ratio of light diffracted into these two orders is controlled by the fine geometry of the grooves.

Traditionally diffraction gratings have been ruled mechanically, a laborious process subject to a host of subtle

inaccuracies, all of which detract from grating efficiency. The standard for mechanically ruled gratings is set by diamond tools and a gold substrate. Fortunately the (negative) master grating can be replicated in resin, much as vinyl records are pressed, to mass-produce gratings which generally exceed the master in efficiency [Bausch & Lomb, 70]. A newer technique for making gratings interferes monochromatic light in photosensitive media to produce so-called holographic optical elements (HOEs). In general these suffer from reduced efficiencies because their groove profile is sinusoidal rather than sawtooth (theory predicts the best performance for a groove apex angle of 90°) [OISPD, 84]. Ion beam etching has shown promise, and a patent is now pending for a method which could set the new standard for optimal, inexpensive, large area gratings [Dave', 87]. This uses multiwavelength interference and a photoresist coating to develop groove profiles of arbitrary accuracy.

Including a Littrow grating in the planetary resonator will serve four purposes. First, since the basic Littrow geometry retrodirects the majority of incident light, such an optical element splits the ring. That is, the oscillating field will no longer be a traveling wave in a ring laser cavity, subject to bidirectional frequency-shifting effects, but rather a standing wave in a linear, but bent, laser cavity. The field will travel "back and forth" as though in a conventional laser that has been wrapped around the planet. Second, the grating will allow oscillation exclusively for one CO_2 spectral line, because all others' frequencies will not survive more than one round trip, getting scattered out of the cavity instead. Third, simply rotating the grating will tune not only to the desired spectral line, but also to the desired μrad point-ahead angle required by the satellites' orbital motion (since as already noted, mere retrodirection would miss the target). Finally, when properly designed the grating will

couple out in a single order all the laser energy not returned to the cavity, uniformly across its extent, yielding a spatially homogeneous output beam.

To settle on a reference grating design we need to know the precise wavelength we wish the resonator to support. Recall from Chapter 2 that the laser bands for the CO₂ molecule are determined by the energies of its vibrational levels, but that the specific lines within those bands are determined by much more finely spaced rotational levels. The distribution of molecules among those allowed rotational levels is independent of their vibrational excitation, depending instead only on the molecular species and its kinetic temperature. That distribution is easily calculated for CO₂ at Venus by the simple program LSRLIN in Appendix A7-6; the result is graphed in Figure 7-4. Collisions ensure that molecules end up in the most populated levels with greater probability, so those levels always have more molecules available for stimulated emission. Although any line near the distribution peak would do, we will choose the highest, $J = 12$. Since the P-branch transition is slightly more probable than the R-branch transition, we choose the P(12) CO₂ laser line for oscillation. For the most common CO₂ isotope (¹⁶O¹²C¹⁶O) that line's frequency is 28.5160266734 THz, a vacuum wavenumber of 951.19226360 cm⁻¹ [Freed et al]. Our design laser wavelength is therefore 10.513 μm.

Appendix A7-7 shows that the comfortably broad efficiency peak for a high-blaze-angle Littrow grating allows us to specify a reference grating with groove spacing $d = 11.68 \mu\text{m}$ for our preliminary design. Detailed design would yield the particular groove profile (deviation from a right apex angle) necessary to commit a few percent of the diffracted energy to the $m = 0$ order for output coupling, and return the rest to the cavity using the $m = 1$ order with a slight angle offset.

We need only note here that such design is possible; we will allow a generous inefficiency loss margin to cover such detuning. The desired groove profile could be photoetched directly on the surface of reflector mirrors and gold-coated, using the Dave' process already outlined. Additional calculations in Appendix A7-8 demonstrate that such a grating has such great spectral dispersion that light from even the closest undesired CO₂ lines cannot remain in the cavity. Tuning to just one CO₂ line is thus assured.

Vertex Station 1

We have designated Vertex Stations 2, 3, 4 and 5 as identical satellites which bend the laser around Venus. Station 1 must therefore be specialized to act as both "ends" of the cavity, one of which uses the grating just described simultaneously to select the emission line and couple out a useful beam. Because large independently-controlled mirrors are required, as well as redirecting hardware to keep the beam aimed at the collinear libration points, our first approach is to configure Station 1 as a co-orbiting constellation of separate satellites, each serving a distinct function. Station 1_a will redirect the beam back to Station 5 and hence on its way back around the ring. The only essential difference between this "end" reflector and the basic Vertex Stations is that, being virtually normal to the beam, it needs only a circular outline. With a smaller area than the others, it will have only about 60 % the number of components (mirror segments, sensors, actuators, and so forth) that they do. Also, 1_a's mirrored surface must fly with a nominal pitch bias of only -36° instead of the planet-facing -90° of the others, which makes it more debris-vulnerable, and requires

particular attention to budgeting moment of inertia as explained earlier.

Station 1 β , the other end of the resonator cavity, will fly just over 1 km away from 1 α , but facing "the other way". Fronted by mirror segments etched completely with the blazed diffraction grating profile designed in the last section, 1 β is the planetary laser's frequency selecting output coupler, scattering almost all incident laser light back to where Station 2 will be when the light arrives there. Nominal operation requires its incidence angle to equal the blaze angle θ (26.75°), so the satellite reflector profile must be elliptical with minor axis 1 km and major axis 1.12 km, giving it about 66 % the number of components as the basic vertex stations. Although similar to 1 α in size, its flight attitude is considerably more peculiar, due to polarization constraints.

The 99.7 % minimum reflectivity expected for the basic vertex reflectors depends on the laser radiation field being s-plane polarized --- that is, with its **E**-vector always normal to the plane of incidence (Figure 7-5). Such reflection from conducting surfaces (metal mirrors) preserves the linear polarization [Jenkins & White, 76], so an s-polarized field can indeed be supported sequentially by satellites 2, 3, 4, 5 and 1 α . The **E**-vector will thus be normal to the satellite orbit plane, which is the plane of incidence. Since the process of stimulated emission preserves polarization as well as phase and frequency (Chapter 2), the gain medium also will maintain s-polarization. Finally, the high efficiency of the tuning grating also requires s-polarization, but for a diffraction grating this means having the **E**-vector normal to the groove pattern (Figure 7-5) [OISPD, 84]. Therefore the grooves of 1 β must be parallel to the orbit plane, requiring in turn the

grating normal to be tipped with respect to the orbit plane by the grating incidence angle of 26.75° .

The 1β satellite attitude is visualized most easily by imagining it first as pitched normal to the incident beam (-144° relative to the velocity vector) like 1α , but then pivoted 26.75° about its long axis in the orbit plane. The order $m = 0$, which is the outcoupled beam, will thus be directed out of the orbit plane (below it, in our reference configuration), making an angle with it equal to twice the blaze angle, or 53.5° . The 1β satellite will therefore fly geometrically "three-axis stabilized planet oriented" because of its pitch and yaw biases; unless inertia properties compensate asymmetrically for this skewed attitude, the resulting large constant gravity gradient torque will exact a continuous logistical propellant penalty.

With the 1β satellite we have finally bounded the planetary resonator cavity; all further elements in the optical path constrain the laser's use, not its genesis. So before investigating the rest of the Station 1 constellation, and the modulators beyond, we should analyze some basic properties of the resonator we have arranged.

Loss Accounting

We start with loss accounting for the cavity, to verify that oscillation can indeed build up. The planetary laser is prey to most common loss mechanisms, as well as peculiar ones of its own. In order, we will examine the distributed losses (imperfect transmission through the medium, diffraction loss and pointing loss) and then the localized losses (reflection

losses, grating inefficiency and coupling loss). The sum of all system losses must not exceed the single-circuit gain (which we have taken to be 10 %) at all times if the cavity is to sustain laser oscillation.

The laser medium comprises only 4.5 % of the cavity path length. At least two of the five tangent regions provide negative loss -- that is, gain -- at all times. The other, non-inverted regions, which we may call dark, are inert to the transiting beam. Being optically thin at the laser wavelength, they cannot contribute appreciably to absorption or scattering of the light, and may be ignored. The other 95.5 % of the cavity path is simply vacuum, incapable of attenuating the beam either.

Diffraction losses, due to finite mirror dimensions, constitute an important loss for typical lasers. The edges of a finite aperture (including a mirror), in diffracting the beam, cause it to spread. A distant target mirror of the same size will therefore intercept, and return to the cavity, only a portion of the beam. The rest, passing by unintercepted, constitutes the diffraction loss per pass. The common trick for reducing such loss is to make the resonator confocal -- by curving the mirrors toward each other, the beam becomes concentrated on the distant mirror and much less is lost off its edges. Our laser, however, has a somewhat different problem.

The width of a laser cavity can support many higher order transverse electromagnetic modes (TEM) than just the simple Gaussian radial power distribution of TEM₀₀. Figure 7-6 represents a few low order modes, showing their transverse phase inhomogeneity. These field amplitude reversals across the beam, if maintained in the far field, would interfere with the heterodyne detection our system presumes, because such a

receiver, beating the signal against a single-mode laser, depends on signal amplitude, not just power. At worst, the receiver might experience signal dropouts as it swept through an inhomogeneous beam pattern. Restricting the planetary laser to the TEM₀₀ mode alone would prevent such interruptions.

Higher order modes might not be a problem in the far field in any case, because such modes suffer more spreading upon exiting the system aperture [Siegman, 86]; their energy should dissipate more, rendering it less detectable at the target than the lowest mode's. This of course indicates that much of the laser's total energy is transmitted in unusable modes, reducing its effective brightness at distance. We can limit the higher order modes anyway, though, if it turns out important to do so. Diffraction losses within the cavity are higher for the higher order modes, too. Figure 7-7 shows the diffraction loss for several of the lowest order modes, plotted against system Fresnel number. With flatter curves, confocal mirrors lose much less by diffraction, even for much smaller Fresnel numbers, than do plane mirrors. Note also that the confocal case allows a more exaggerated loss dispersion (hence easy selection) among the transverse modes for a given Fresnel number.

For our use, take the medium index of refraction $n = 1$, wavelength $\lambda = 10.5 \mu\text{m}$, and the mirror spacing $l = 8983 \text{ km}$, the line of sight separation between vertex stations. Then with a 1 km beam (radius $a = 500 \text{ m}$) and even with plane mirrors, the Fresnel number at 2650 is so far off the top of the graph that diffraction losses can be completely neglected, at least for many low order modes. Thus the cavity will distribute laser power in all of them, possibly with negative far-field consequences as discussed above. One solution is to allow an arbitrary TEM₀₀ total cavity diffraction loss of 1% (budgeting $1/5$ to each ring sector) limiting our beam

diameter to only about 100 m. Even the very next transverse mode would then suffer about 2.5 % total diffraction loss, perhaps enough to keep it from oscillating. But reducing the beam diameter by 10 means reducing the laser communication capacity by 10^4 , a drastic penalty. An alternative solution seems available, though.

Just as the flat confocal resonator curves lie below the plane mirror curves in Figure 7-7, yet steeper curve families for various unstable resonators (mirrors convex toward each other) should lie above those for the plane case. Therefore it would be possible to select a subtly convex mirror curvature such that our system Fresnel number of 2650 (corresponding to the 1 km beam diameter) introduced acceptable diffraction loss into the lowest order transverse mode and unacceptable losses into all higher, undesirable modes. By operating "on the edge" of controlled diffraction loss, then, the system could prevent unwanted mode oscillation and thereby ensure spatial coherence across the beam. That edge is not as sharp as we might assume by noticing that the loss dispersion among modes decreases as the curves get steeper, because a special property of unstable resonators is that the diffraction loss difference between the lowest mode and all the others is peculiarly great (Svelto, 84), permitting "good transverse mode discrimination".

The curvature required would be so slight as to be macroscopically invisible. It could never be nearly so large as a reverse confocal curvature (all the rays would walk out of the cavity!), and yet Appendix A7-9 shows even that extreme to result in a center-to-edge deviation from plane of only 1.4 cm for mirrors 1 km across and separated by the 8983 km line of sight. Any resonator mirror curvature called for by more detailed optical design would therefore represent a minor alteration of our reference plane design. In fact, active

mirror control would be able to effect substantial desired convex or concave figure adjustment as easily as it could maintain plane reflectors. We will specify plane mirror surfaces, but budget 0.2 % diffraction loss per ring sector.

Pointing loss, resulting from mis-aiming the mirrors at each other, is a loss variety unique to orbital lasers like ours. The resonator components of conventional lasers, being mounted on a common, stable metering armature, can be aimed optimally by adjusting their positions and orientations until the measured laser output peaks. Our system will do the same, but we cannot then lock the components into position. Since they move through space individually, their beam alignment will constitute a continual active task, particularly demanding since in general each satellite must preserve the reflected line of sight between both its neighbors. The result will be limited by additive errors due to measurement inaccuracies and controller resolution, collectively called pointing accuracy. Chapter 8 discusses how we can achieve accurate pointing, but here we acknowledge its imperfection and specify its performance by budgeting another 0.2 % loss per sector. Appendix A7-10 translates this loss allotment to a pointing accuracy requirement for each satellite of 0.22 μ rad, almost 7 times less stringent than the Hubble Space Telescope specification.

Reflection loss at the mirrors is our first localized loss type. We have seen that s-polarized 10.5 μ m light reflecting at incidence angles greater than 45° from an enhanced gold coating over a microsmooth substrate suffers less than 0.3 % loss due to scattering and absorption, so this value applies to Vertex Stations 2, 3, 4 and 5. Satellite 1 α , encountering the beam at virtually normal incidence, introduces the slightly higher reflection loss value of 0.5 %. Satellite 1 β 's diffraction grating also contributes reflection loss; the

blazed orientation means essentially normal incidence (see Figure 7-3), and we may assume the grating to be bare gold (for maximum efficiency). Thus the loss value for unenhanced gold (0.6 %) should be used.

Without a specifically detailed, tested design, the diffraction grating's absolute efficiency cannot be known. The development of Appendix A7-7 indicates that only fabrication and operation imperfections would cause unwanted scattering loss from the grating. Allotting a 2 % loss for grating inefficiency is probably excessive, but our necessary uncertainty warrants conservatism.

Finally the 1β grating must also scatter some laser light into the $m = 0$ order as coupling loss, if we are to have any output. In operational fact, the coupling loss will be whatever is left over after all other single-pass losses are subtracted from single-pass gain, but we should have some idea of the coupled fraction magnitude, since it would partially determine a detailed grating specification. Typically a few percent of the circulating power gets coupled out of working lasers. We assume that in steady-state operation the planetary laser will extract and emit power at the solar pumping rate, so that the percentage complement to the pump-limited coupling loss must be the circulating power. We will specify coupling somewhat arbitrarily at 2 %.

Now we can verify that system gain exceeds all losses, satisfying that vital requirement for oscillation. Figure 7-8 shows that we must choose the orbital position of our split ring carefully for loss accounting. The 1β end of the ring contributes the greatest losses, so the worst case occurs when the tangent regions closest to that end are dark; then the light must survive yet more losses before reaching its next dayside replenishment passage. Figure 7-9 illustrates in

detail the gains and losses which accumulate during one round trip when the ring is in this worst-case position. Because all the values we have assumed are small percentages, we simply add them; though not rigorously accurate, the error is less than the defensible accuracy of our assumptions anyway. For simplicity the pentagon is oriented such that two dayside gain regions contribute equally; each is allotted half of the single-circuit 10 % gain. We start arbitrarily at 1α with nothing, then add laser gain and subtract pointing, diffraction and reflection losses all the way around to 1β . There we subtract coupling, grating and reflection losses before returning back around the ring to 1α . The critical number in this diagram is the minimum, which occurs after the return pass at Station 4, immediately prior to replenishment, and must exceed zero for laser oscillation to build up. The 0.5 % margin at that point provides a measure of how much total error would be allowed our assumptions (including the Venusian mesospheric gain value) before lasing became impossible with this resonator.

Cavity Control

Knowing that the laser amplification condition can be met, we must now see what kind of control the cavity components require if the laser is to operate. The first and simplest adjustment is the slight point-ahead bias already mentioned as necessary to maintain a continuous optical path in the rotating ring. The fact that light takes .03 s to travel between vertices perhaps more clearly than any other measure reveals the sheer size of a planetary laser. Calculated and listed in Appendix A7-11, these mirror pointing biases differ among the vertex stations because the split ring is not quite a regular

pentagon. The smallest of them is comfortably 10 times the pointing error budgeted in the last section. Except where these small angles affect spacecraft operation, we will not refer to them in the remainder of this study.

Recall from Chapter 2 that laser light, in addition to being "monochromatic", is also coherent. All locations in the radiation field across the beam thickness will share the same phase (spatial coherence), and any two points along the field's length will bear the same phase difference at all times (temporal coherence). In the last section we saw how restricting transverse modes allows spatial coherence. Next we see that temporal coherence and monochromaticity are related, and evolve together the most challenging system performance constraint for the planetary laser.

The distance over which light remains temporally coherent is called quite naturally its coherence length. The literature on laser applications is rife with comments about the "essentially unlimited" coherence length of laser light. While this hyperbole is understandable coming from grateful researchers whose early work antedated the seemingly miraculous abilities of lasers, we must ask just how unlimited "essentially" means when engineering a planetary-scale laser. The coherence length of the interstellar beam itself is not particularly relevant (heterodyne detection does not make use of the phase properties of the signal, as for instance interferometry does) --- in fact it is moot --- but the intracavity coherence length is crucial. That is, if the field loses its temporal coherence before repeating itself (if to first order the light's coherence length is shorter than the cavity itself), we will not have a laser.

Coherence length is simply the length traveled by light during its coherence time τ , which in turn is simply the

inverse of its emission linewidth $\Delta\nu$ [Jenkins & White, 76]. Thus the greater a laser's spectral purity, the longer is its coherence length. What determines its spectral purity? The "natural laser linewidth" of CO₂ is dominated by broadening effects as noted in Chapter 2, primarily by homogeneous (pressure) broadening above 10 torr and by inhomogeneous (Doppler) broadening below that [Siegman, 86]. The atmospheric pressure at altitude for Venus may be taken as six orders of magnitude smaller [Deming & Mumma, 83] than this transition, so Doppler broadening will dominate. Appendix A7-12 calculates the half-peak natural width to be 43 MHz.

But this represents merely the "gain profile", an envelope in frequency space within which a cavity mode must fall if lasing is to occur. The laser's actual oscillation frequency will be pulled slightly towards the natural line center, and will be gain-narrowed to a possibly minute fraction of the gain profile width. The lineshape function describing the gain profile is Lorentzian, with the property that its integral over the frequency domain is unity. Thus as it intensifies, its width must decrease. With repeated oscillation in the cavity, then, the lineshape function multiplies itself, becoming eventually like the Dirac delta "function". Statements can be found in the literature that this gain-narrowed output linewidth is "infinitely" narrow, but again, we must quantify just how narrow this is.

The ultimate "limit to monochromaticity" is set by spontaneous emission noise in the medium, and can be calculated from quantum mechanical theory. Typically, that limit is of no practical value (Appendix A7-12 shows it to be $3(10^{-26})$ Hz in our case). Frequency variations due to cavity length instability completely dwarf the quantum noise fluctuations in working lasers, setting the real limit to their monochromaticity [Svelto, 84]. In the laboratory and in the

field, temperature variations and mechanical vibrations of the cavity armature are the primary causes of "the short-term frequency jitter and the long-term frequency drift" of laser oscillators [Siegman, 86]. Under special actively-controlled circumstances, some lasers can achieve spectral purities of 1 part in 10^{13} (which would correspond to 3 Hz for our laser), making them the most accurate clocks known.

Understanding that our ability to control the planetary laser's cavity length will limit its monochromaticity, which in turn will directly limit its coherence length, enables us to develop quite simply the strictest performance specification of this entire study -- one which will completely drive the design of supporting subsystems. Requiring the laser coherence length to match the pentagonal perimeter round trip sets a lower bound on coherence time, which then determines the maximum allowable oscillation linewidth. Finally, that frequency variation limits the uncontrolled cavity length variations permissible during operation. The sobering result, also from Appendix A7-12, is that if the uncompensated resonator cavity length changes exceed $17.3 \mu\text{m/s}$, it will not lase.

Although developing an approach to meeting this requirement will consume Chapter 8, we should recognize here as preview that the space environment helps. The perfect physical component isolation which only free-fall allows, immediately shifts the cavity stability problem to one of spacecraft control. We should also note that, tough as it will be to meet the 62 nm and $2 \mu\text{m/s}$ mirror-positioning requirements from Appendix A7-12, and challenging though it will be to control such motions cooperatively over planetary distances, achieving these successfully will automatically ensure an output beam of such spectral narrowness that far-field detectors could track its long-term frequency drift using extraordinarily narrowband channels. The resulting superb signal-to-noise ratio would

thus enable enormous data transfer rates. If the planetary laser can be made to work at all, then, it cannot avoid being an excellent interstellar communication device.

The Output Beam

Based on the photon volume emission rate of the natural Venusian laser [Deming & Mumma, 83], Appendix A7-13 calculates our circulating cavity power to be 8.7 MW. Distributed over the large cross section of the beam, this yields a power density of only 11 W/m^2 , less than 1 % of the solar intensity at 1 AU.

The beam coupled out by the 1 β grating contains 180 kW of power. Three considerations make it desirable to reduce the diameter of this output laser beam. First, the final system aperture should be much smaller than 1 km anyway. Although diffraction will cause the interstellar beam to spread, the 1 km cavity aperture is so large compared to IR wavelengths that the emergent full-size beam would diverge to a spot size only 1 % the size of Mercury's orbit 4.3 ly away, instead of covering a reasonable habitable zone. Were we to rely on diffraction spreading alone to cover distant planetary orbits, Figure 7-10 (plotted by the program APERTURE in Appendix A7-14) shows that system aperture diameters between roughly 1 m and 200 m would be appropriate. So reducing the beam size immediately after it is coupled out of the cavity begins to prepare it for the interstellar trip.

Another advantage to reducing the beam size arises because, if a large duty factor is to be maintained, Station 1 has to switch the output beam from one libration point

modulator to the other twice per orbit. The large (approximately 180°) direction change implies that mechanical methods be used; quite apart from accuracy considerations, minimizing downtime during switching favors as small an opto-mechanical device as possible. Slewing a 1 km mirror quickly, smoothly and accurately is a challenge we would rather avoid. Finally, a narrower beam allows smaller modulator optics.

Beam reduction cannot be excessive, however. Although the total beam power is not unconscionably high, its power density becomes quite high as its area is reduced. If we continue to assume 99.5 % reflectivity for enhanced gold reflecting optics at near-normal incidence angles, and if we assume the other 0.5 % to be absorbed by the mirror (pessimistic, since some portion of that amount is scattered rather than absorbed), the mirror heat load can be plotted as a function of mirror diameter for our laser. Figure 7-11 shows the drastic thermal loads encountered by small mirrors (Appendix A7-15 gives the program). We choose a reference beam diameter -- called the intermediate beam -- of 10 m. This represents a 10^4 area reduction, putting the beam optics just before the knee of Figure 7-11 with a worst-case absorbed power density of 11.5 W/m^2 , again only about 1 % of the solar flux at 1 AU and therefore matched to the intracavity condition. The worst-case total heat load will of course be 900 W, which must be removed, probably by simple radiation, without optically disturbing vibrations (as from transfer fluid turbulence).

Beam reduction is performed by the 1Y satellite of the Station 1 constellation. Similar in general to the other large reflector satellites, its mirror segments are both uniquely figured (fabricated) and configured (operated actively) to form an off-axis symmetrical paraboloid of parameter $p \approx 1.40 \text{ km}$, capable of focusing the 1 km beam to a point. As explained

earlier, 1β scatters the output beam below the nominal orbit plane; 1γ must therefore fly three-axis planet-oriented, with its geometrical center approximately 0.4 km below the orbit and 1.0 km south of the orbit plane. The formation-flying requirement means that 1γ 's slightly non-Keplerian orbit (non-planet-centered, and moving at too slow a velocity for its altitude) must be maintained either propulsively or structurally (Chapter 8).

To avoid those kinds of secular effects for the rest of the Station 1 constellation, 1γ is oriented to focus the intermediate beam back up into the orbit plane (this orientation elliptically polarizes the beam), to a point roughly 0.8 km below the orbit altitude but in radial line with the pentagonal vertex. There an inverse 15 m off-axis symmetric paraboloid mirror (the 1δ satellite) with parameter $p \approx 10$ m intercepts the beam before it comes to a point focus, almost completely "undoing" 1γ 's convergence by recollimating the laser light into the just slightly converging 10 m intermediate beam, and redirecting it radially outside the ring through the gap between 1α and 1β .

About 1 km above the orbit altitude the intermediate beam is intercepted once again, by the wheel-shaped 1ϵ satellite. Tethered to 1δ 1.8 km below it, 1ϵ serves as an optical switch to send the beam on to one of the two modulators. Its mechanism consists essentially of a 100 m controlled structural ring supporting two diametrically-located, separately pivoting 15 m plane mirrors. By coordinating these mirrors' rotations, rotation of the entire 1ϵ wheel, and slight oscillation of 1δ to track the mirrors as 1ϵ rotates (Figure 7-12), the system can sustain line-of-sight contact with one or the other of Venus' collinear libration points virtually all the time.

The reflection geometry just outlined may at first appear more complicated than necessary, but several factors favor it. Because the libration points lie in the resonator orbit plane, the redirecting switch need only change its attitude about one axis (normal to the orbit plane) to sweep the beam in that plane, if it receives the beam also from within that plane. Since $l\gamma$ must be out of the orbit plane, using an additional reflection -- $l\delta$ -- greatly simplifies operation of the $l\epsilon$ switch. Also, as $l\epsilon$ rotates to keep aiming the beam at a libration point, $l\delta$ must oscillate by an angular amount equal to the arc subtended by the separation of $l\epsilon$'s rim mirrors, 56 mrad. Such oscillation is much more feasible using a small satellite like $l\delta$ than it would be using $l\gamma$ directly. Finally, $l\epsilon$ must be well outside the ring so that $l\alpha$ and $l\beta$ do not obstruct its view of the libration points when Station 1 overflies the terminator. By tethering $l\delta$ and $l\epsilon$ together directly in the ring's gap, careful mass management can insure that they both remain in formation with the rest of Station 1 without continual energy expenditure; furthermore, the pair flies in a principal-axis, stable gravity-gradient orientation, so no attitude torques will accumulate.

$l\delta$ is small enough, at 15 m, that its rather noticeable paraboloidal curvature should be composed of smaller segments than those we have discussed for the larger satellites. Each piece must be unique, and making them on the order of 1 m across instead of 3 m would probably simplify their manufacture. On the other hand, the pivoting mirrors of $l\epsilon$, also 15 m in total diameter each (to accommodate the 10 m beam at non-normal incidence), might profitably be made monolithically diffraction limited. They would of course incur a mass penalty commensurate with maintaining their stiffness to the $\lambda/20$ surface figure tolerance, and their fundamental vibration frequency would drop to about 120 Hz, but controlling the quick large rotation of such passive plates

would be much simpler than the similar rotation of a complex, active segmented plate with its support structure. Electromagnetic "space bearings" (Chapter 8) could be used to mount, isolate and move them, completely obviating any physical contact between them and the 1ϵ spacecraft bus. Since both these 1δ and 1ϵ optical element specializations represent merely refinements of the basic mirror technology already discussed in detail, we will not concentrate on them any further.

The pointing accuracy required of these latter Station 1 satellites among each other is trivial compared to the intracavity requirement already specified (and addressed in Chapter 8), because they are so close together; consequently we need not work with it in detail. 1ϵ 's pointing accuracy to the libration points turns out to be the most challenging in the entire system, however. We have already noted the slight convergence applied to the 10 m intermediate beam. The reason for this is that a 10 m plane wave emanating from the 1ϵ wheel would diverge to a spot diameter of 2600 m at the Venusian L2 distance; a 10 m optic at L2 would intercept only $1.5(10^{-5})$ of the central Airy spot, or $1.2(10^{-5})$ of the beam's energy (Appendix A7-16)! Instead we let the intermediate beam continue converging slightly after its reflection from 1δ , so as to produce a spot only twice the projected diameter of the L2 optics, or 20 m. While this allows collecting a full 21 % of the beam's energy for modulation and stellar targeting, it does impose a severe pointing accuracy constraint on the 1ϵ optical switch. Appendix A7-16 goes on to show that covering the presumed L2 optic with such a spot requires 1ϵ to have a pointing accuracy of 5 nrad, which is about seven times as good as the Hubble Space Telescope can achieve.

The difficulty does not consist so much of knowing where the target mirror is (acquisition), because unlike an inherently uncooperative astronomical target, the L2 station will be in round-trip feedback communication with \mathcal{L} at 0.3 Hz; rather, the problem is one of holding aim on the target (attitude stability) for long times as the \mathcal{L} wheel rotates and its mirrors pivot. In Chapters 8 and 9 we investigate methods for meeting the 5 nrad specification for the \mathcal{L} mechanism. Fortunately this error, introduced due to the huge absolute separation between Venus and its libration points, can be cancelled by the pointing system of the modulators so that it does not contribute to the interstellar pointing error. That is, although \mathcal{L} cannot hold its distant spot exactly centered over the L2 receiving optic, the energy spilling over the edges can be used to advantage. A collar of sensors around its primary optic will enable the L2 station, through knowing within centimeters the location of its received beam pattern, to derive at any time the actual \mathcal{L} pointing error to μ rad resolution (called post-determination). That error can then be applied as a tilt correction in the L2 output train, preventing it from adding to the interstellar pointing error. Similar overspill sensors will enable \mathcal{L} itself to decouple its pointing performance from the accumulated micro-errors of the earlier optical path.

Libration Point Modulator Stations

L1 Station and L2 Station, the two alternative final system components of our interstellar transmitter, perform the same job from opposite sides of the planet. In facing Venus, L2 Station also faces the sun. Since L2 is only about $60(10^3)$ km beyond the apex of Venus' umbral shadow, the sun

will appear to L2 Station as a ring 500 μ rad thick with radius 6.2 mrad, being centrally eclipsed by Venus. Thus the sensors and telemetry receivers that L2 Station uses to maintain contact with the resonator ring will have 15.5 % of the bright and noisy sun in their field of view. (The high signal-to-noise ratio required to limit bit errors despite this 400 W/m² of broad spectral noise provides one of several strong reasons to use laser links for interstation telemetry.) Apart from such effects due to their differing solar orientations, though, the two libration point stations can be considered to be identical; we will therefore continue to refer only to L2 Station.

Each will use a double-reflection geometry to enable full sky coverage. The fixed location of L2 relative to the Venus-sun gravitational system becomes a great advantage; once locked onto a particular star, the greatest angular excursions necessary are those to compensate for Station 1's orbital motion around Venus (15 mrad/hr), and Venus' orbital motion around the sun (5.3 μ rad/112 d even for the nearest star). Such compensation, as well as the angular error correction already mentioned, will be performed immediately by the receiving mirror, a three-axis stabilized 15 m reflector satellite called the Transducer.

In our reference configuration, the Transducer also modulates the laser beam and aims it at a given target star system, shaping the wavefront so the beam spread matches that target. Thus for targets on the celestial hemisphere behind Venus, the Transducer can receive, condition, and transmit the beam alone. Targets in the other half of the sky will require one additional reflection by the Ring, an annular satellite 100 m across centered on and normal to the Venus-sun line. By rotating in its plane, the Ring will be able to maneuver its 15 m plane mirror around the intermediate beam axis to avoid

having the Transducer eclipse its target star. Changing the axial separation between the Transducer and the Ring in turn avoids having the Ring eclipse targets seen directly by the Transducer. This simple geometry assures months-long continual coverage of any star in the sky, interrupted only momentarily by mirror switching.

The acceptable final pointing error for the system, which must be controlled by the Transducer (and the Ring if in use), consists of two completely unrelated contributions: the uncertainty with which we can point, and the "designed defocusing" necessary to cover our uncertainty of where to point. The latter component is comprised in turn of two parts: uncertainty about a star system's location, and uncertainty about the target's location within that system. Although in Chapters 8 and 9 we see how to make 1ϵ 's pointing accuracy almost an order of magnitude better than the current Space Telescope standard, for comparison we will accept that standard (34 nrad) for our final stage.

In the years it takes light to reach another star, that star moves. For targeting purposes the relevant component of its motion is across the celestial sphere, transverse to our line of sight; this is called proper motion, and due to parallax is obviously greatest for just those stars of interest to us: the closest. Clearly our signal must be "pointed ahead" to where the star will be when the signal reaches it, a maneuver limited by our ability to predict proper motion accurately. On the timescale of interest for our communication range (roughly 4 to 80 years), proper motion consistently follows a straight line, and estimates of the rate of that motion are made by comparing observations taken over many years. Relative proper motion is measured against the background of "fixed" distant stars; the data approximate absolute proper motion as the motions of farther (and

necessarily fainter) stars are measured. The extreme stability of ground-based observatories on the airless Moon will permit great strides in such measurement.

However, to point our spacecraft we simply need good relative proper motion data, since the inertial-reference star sensors use bright nearby stars themselves. The Space Telescope pointing system is only as good as its star charts, after all, which are currently being updated to 3 nrad accuracy [Gatewood, 87]. Relative proper motion uncertainty μ_p is currently catalogued for many nearby stars at about 24 nrad/yr [Allen, 73]. One exciting instrument being developed for space station use is the Astrometric Telescope Facility (ATF), which as its name implies will measure star positions. Over a 30 yr period, even this contemporary observatory will yield proper motion data for any of our target stars good to $\mu_p \approx 240$ prad/yr [Gatewood, 87], a value certain to improve substantially over the next century with yet better equipment. The corresponding total uncertainty for our farthest stars, 82 ly away, represents the pointing uncertainty we have to build into the laser's operation (by way of beam spread) to be sure of hitting those stars; at 20 nrad, it is of the same order but still 40 % smaller than our (ST referenced) ability to hold the aim anyway.

The largest contribution to our system pointing uncertainty is again one which we must build in purposely: beam divergence sufficient to cover a spot in the far field sure to include a reasonable target within a star system, as discussed earlier for Figure 7-10. Taking current observations of diverse organic molecules in Titan's atmosphere [Sagan et al, 87] as evidence that Saturn's orbit might be an appropriate outer limit for a stellar habitable zone, and taking Mercury's orbit as a sensible inner limit, we can easily calculate the beam spread required to cover this range of planetary orbits

for stars out to 85 ly with the diffracted beam's central Airy spot. The program ANGERR in Appendix A7-17 does this, and Figure 7-13 shows these curves plotted together with the pointing error and proper motion uncertainty we have just discussed. At the worst extreme of our design range (targeting a small orbit about a distant star), the required beam spread due to target size still exceeds our pointing error by a factor of 3. For less extreme orbit choices, and particularly for closer stars, necessary defocusing exceeds even conservative pointing error and proper motion uncertainty assumptions by orders of magnitude.

The decision about how much to diverge the beam intentionally depends on the laser's intended use, and dictates its modulation rate. Covering a generous orbit range with a SETI beacon would dictate quite low data transfer rates (Chapter 4), probably appropriate for first contact purposes anyway. Later CETI transmissions, or those to human stellar outposts or probes, could take advantage of known small orbits close to the star, or even well-characterized receiver orbital parameters, to aim at smaller areas and thus attain tremendous data transfer rates (Chapters 4 and 12).

It should not surprise us that modulation rates differing by orders of magnitude require different technologies to accomplish them, so that details of the L2 Station Transducer, for instance, must vary widely with system intentions. There is therefore no such thing as a generic modulator for the interstellar laser. Furthermore, contemporary modulated lasers are small (fiberoptic transducers and disc readers, for example) while contemporary big lasers typically are not modulated (industrial and SDI units emphasize power, not subtlety). In attempting to devise schemes for modulating the planetary laser, then, we are faced simultaneously with a dearth of relevant published information and a design range

from DC to multi-GHz. However, by pointing out available methods, we can project the starting directions research would necessarily take in developing various modulators for such a large communication device.

Typically amplitude modulation of an optical beam would deflect rather than extinguish it, because any beam power which is not transmitted must be absorbed, transported and rejected by the modulator, an unwelcome if unnecessary thermal burden. Also, we might easily (and correctly) imagine that deflecting a beam slightly would require much less sheer modulating power than neutralizing it altogether. For a tightly collimated beam aimed at a small target, particularly at great distances such as ours, extremely small deflections are necessary to yield quite high far-field extinction ratios. For instance, a beam diverged to cover a Saturn-sized orbit in the very next star system (our widest design beam spread) still only needs to be diverted $70 \mu\text{rad}$ for its central Airy spot to miss the target entirely.

The deflection methods fall into two broad classes, each appropriate for a different modulation regime: low rates are handled best by acoustic techniques, while electro-optic phenomena can attain high rates [Gordon & Cohen, 65]. The simplest acoustic scheme we can propose has a number of merits peculiar to a SETI laser. Imagine a thin mirror plate of stiff, microsmooth metal, gold-coated in the usual way and simply supported around its edge by a comparatively rigid frame (Figure 7-14). An electrically operated actuator (electromagnetic or piezoelectric) positioned behind the plate could control quite precisely its out-of-plane deflection. The proper convexity bias would diverge a reflected beam to cover the chosen target; for example, Appendix A7-18 shows that a $6 \mu\text{m}$ deflection would be proper for targeting an Earth-sized orbit at 10 ly using a 15 m

mirror. A slowly periodic bias variation could remove the planetary resonator gain envelope ripples (Chapter 5) for the chosen far-field distance. Then oscillating the convexity about that bias function would change the beam's spread and hence its dilution, revealed in the far field as amplitude modulation. The obvious tradeoff would be that smaller plates could be vibrated faster with lower power, but would absorb more heat per unit area from the beam (degrading actuator longevity).

Another acoustic method is based on the direct interaction between light and sound, as in SBS. An acoustic wave launched into a proper transparent medium will alter the medium's refractive properties, setting up a volume diffraction grating which scatters the light predictably [Cohen & Gordon, 64]. Refinements using surface acoustic waves can efficiently deflect well-guided optical beams even up to GHz rates, but require microscopic electrodes throughout the medium to stimulate the waves, and require high power [Liu, 86].

Long ago Cohen & Gordon [64] demonstrated that the electro-optic analog to acoustic modulation was capable of 10 GHz rates (microwave frequencies), and this has become the standard method of deflecting small optical beams. Modest powers can be used to remove completely the zero-order beam energy (the undeflected direction), following the modulation envelope of the microwave "subcarrier". Even early attention focused on ferroelectric crystals with high microwave dielectric constants (like $\text{KTa}_x\text{Nb}_{1-x}\text{O}_3$, called KTN, operated near its Curie phase transition temperature of 283 K). A beam traveling through the crystal at right angles to an injected microwave signal will scatter off the induced grating primarily into the first order, under proper conditions. The maximum modulation frequency is limited by the microwave propagation speed through the medium, so large areas would require a

network of microwave sources. Devices using surface electromagnetic waves and optimized slow wave structures are expected to extend to larger sizes multi-GHz modulation rates [Liu, 86].

In connection with these methods operating at microwave rates, we should recognize that control circuitry to support such high data transfer rates is already available. GaAs chips in particular can operate to tens of GHz [Nathanson et al, 87], and (superconducting) Josephson junction digital integrated circuits have been demonstrated with logic gate operation at 100 GHz [Sone et al, 85].

Realistically, without at least a dissertation's worth of original, detailed study, a set of defensible modulators to serve the design range of the interstellar laser cannot be rigorously specified. However, recognizing that emerging modulator technologies can be extended to sizes appropriate for interstellar lasers suffices for this feasibility analysis. Even without new developments, combinations of the acoustic and electro-optic methods presented here can meet the modulation and beam spreading requirements of the interstellar laser's final transmitter stage. Since alternative concepts would completely determine detailed systems designs for the Transducer, we will not consider extensively that satellite's subsystems.

Having developed a consistent set of performance specifications for all the reflecting hardware required to form, sustain, control, tap, condition, modulate and wield the Venusian interstellar communication laser, we must next evolve a fleet of buildable spacecraft capable of realizing the optical path's potential by supporting and servicing these mirrors.

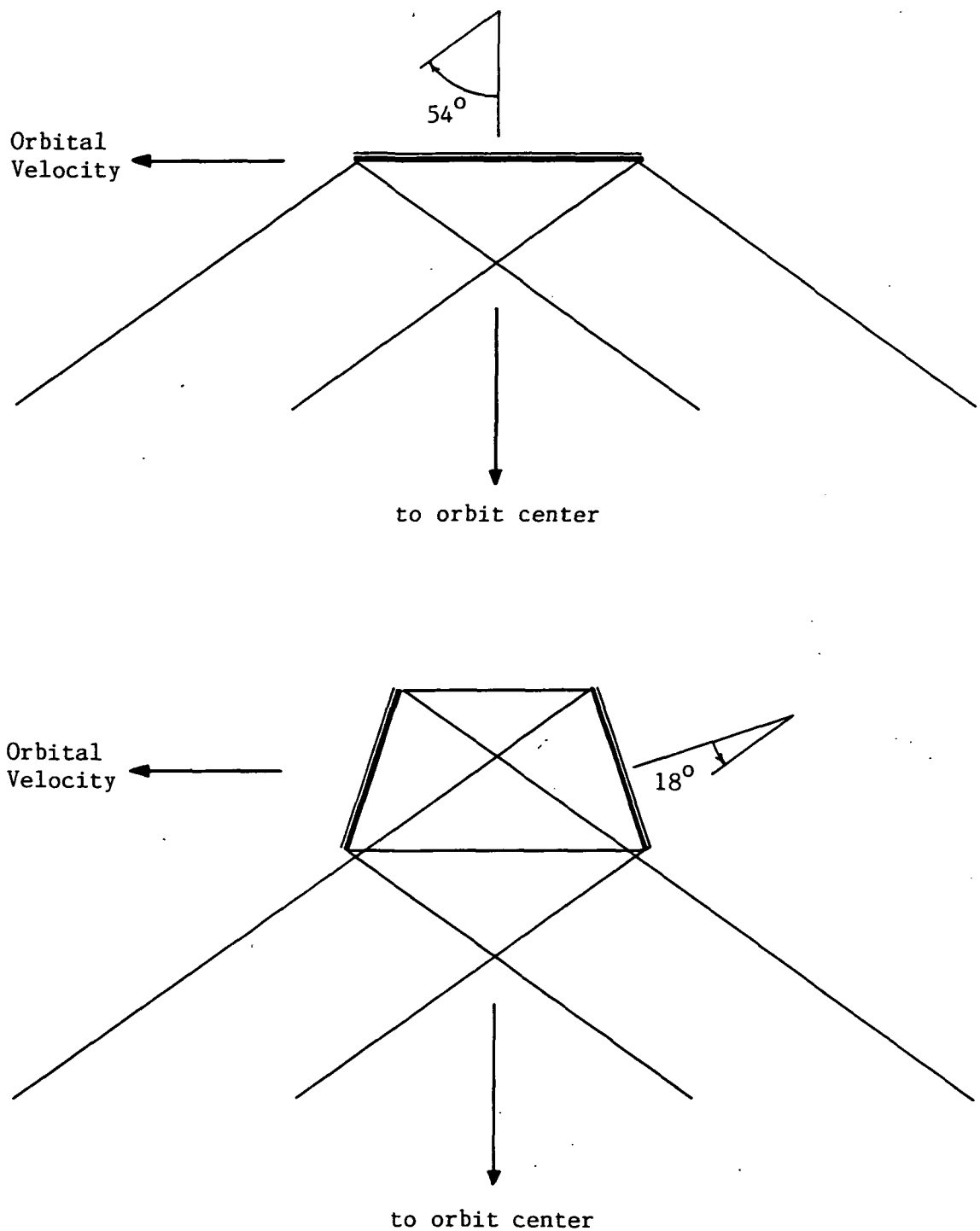


Figure 7-1 At a vertex station, the beam can be folded using a single-reflection (top) or a double-reflection (bottom) geometry.

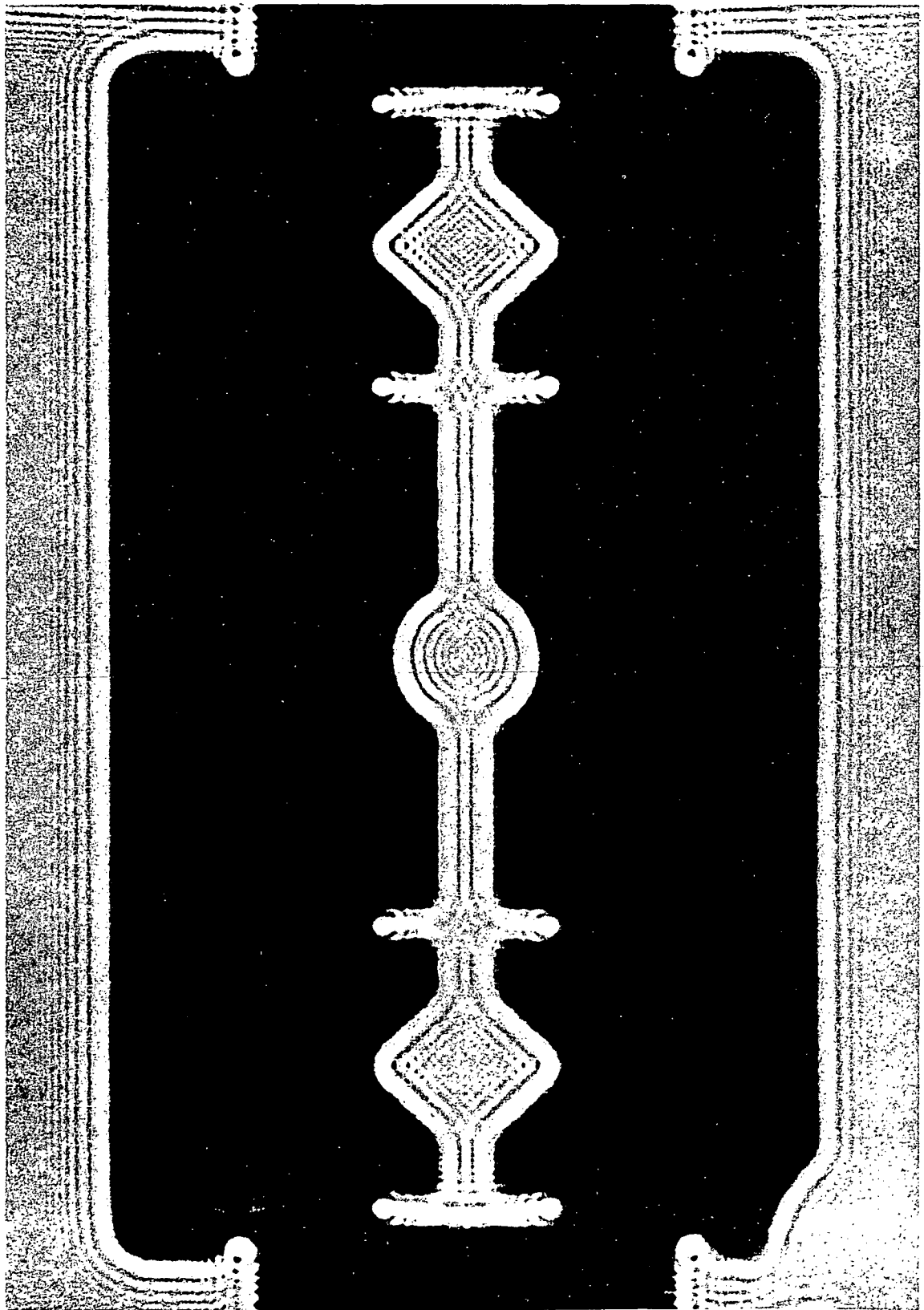
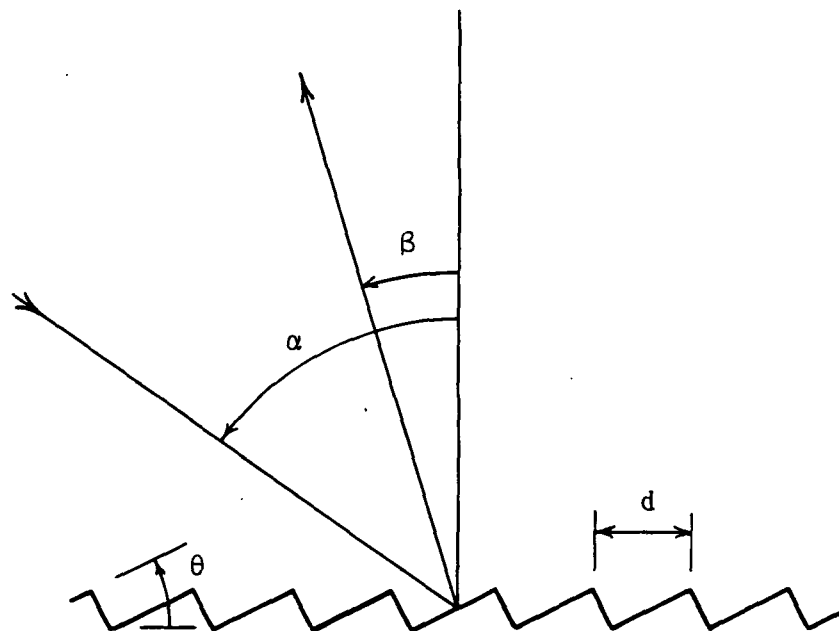


Figure 7-2 Diffraction fringes produced by edges. [Unref, 73]



$\alpha \equiv$ incidence angle

$\beta \equiv$ diffraction angle (< 0 if the diffracted ray is on the opposite side of the grating normal from the incident ray)

$\theta \equiv$ blaze angle

$d \equiv$ groove spacing

$m \equiv$ diffracted order (≤ 0 for $\beta < 0$)

Figure 7-3 Standard diffraction grating nomenclature.

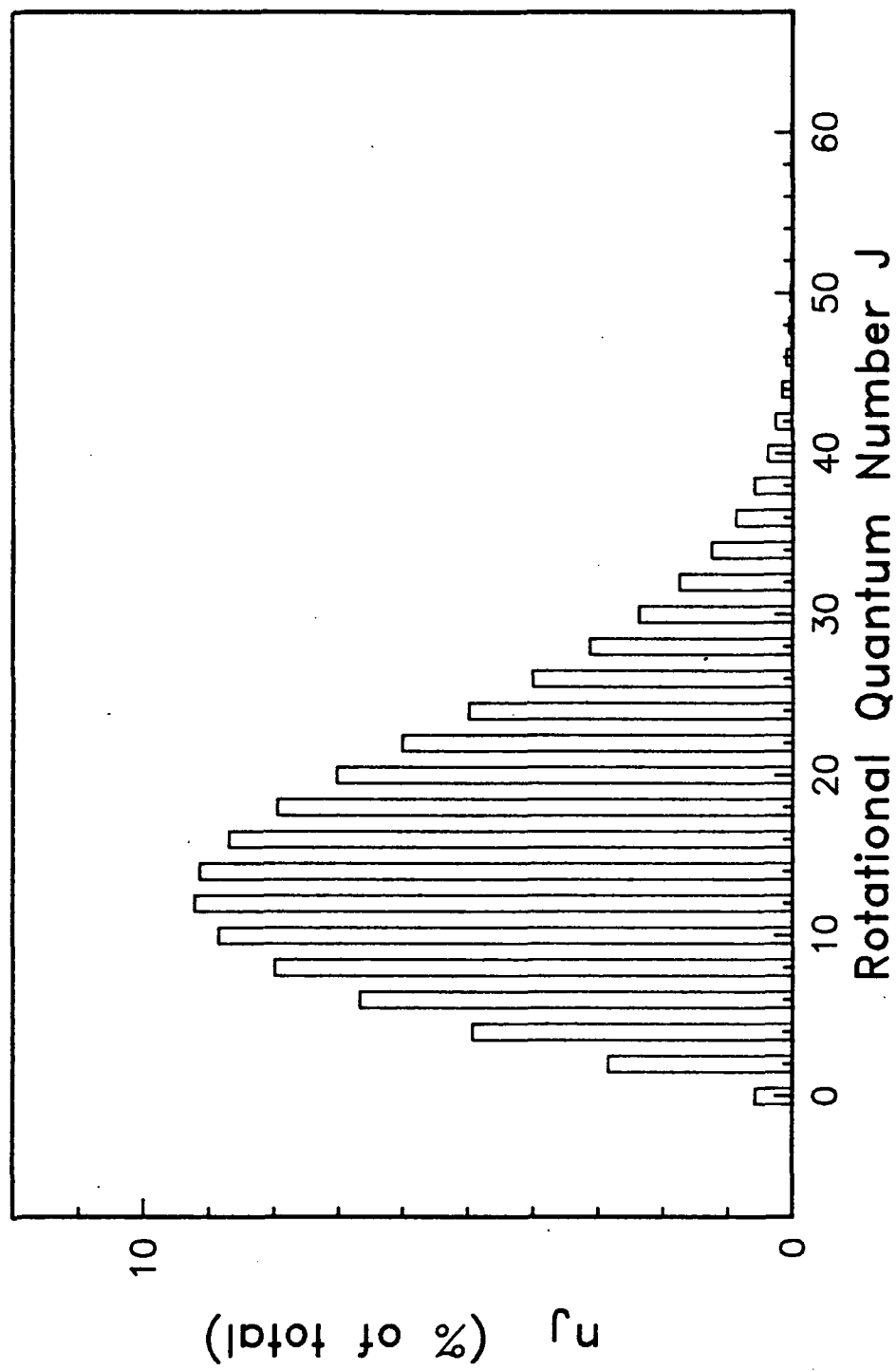


Figure 7-4 Distribution of CO₂ rotational energy levels in the Venusian mesosphere.

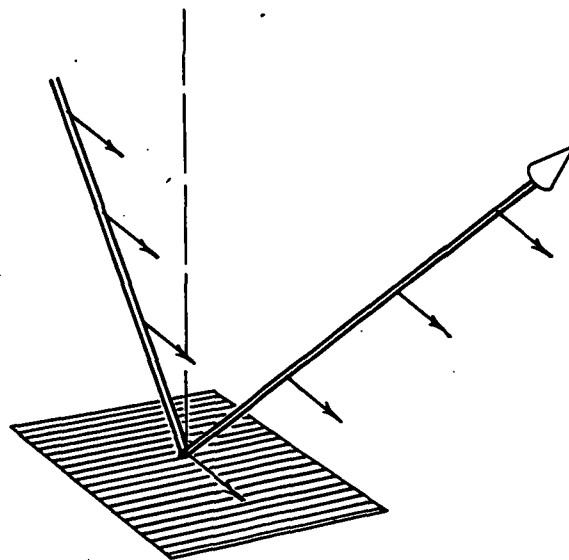
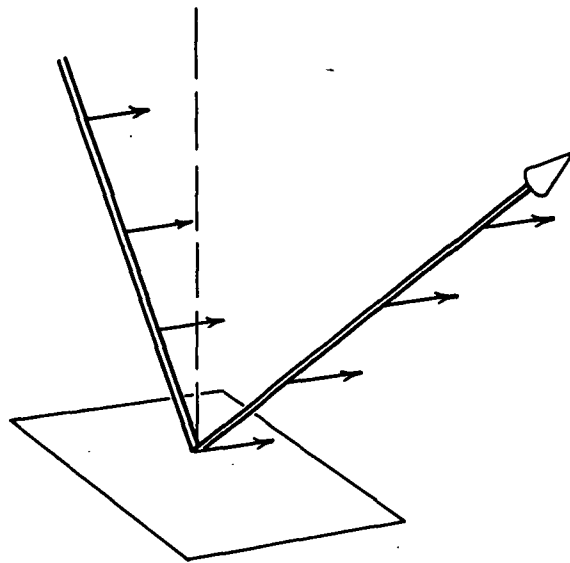
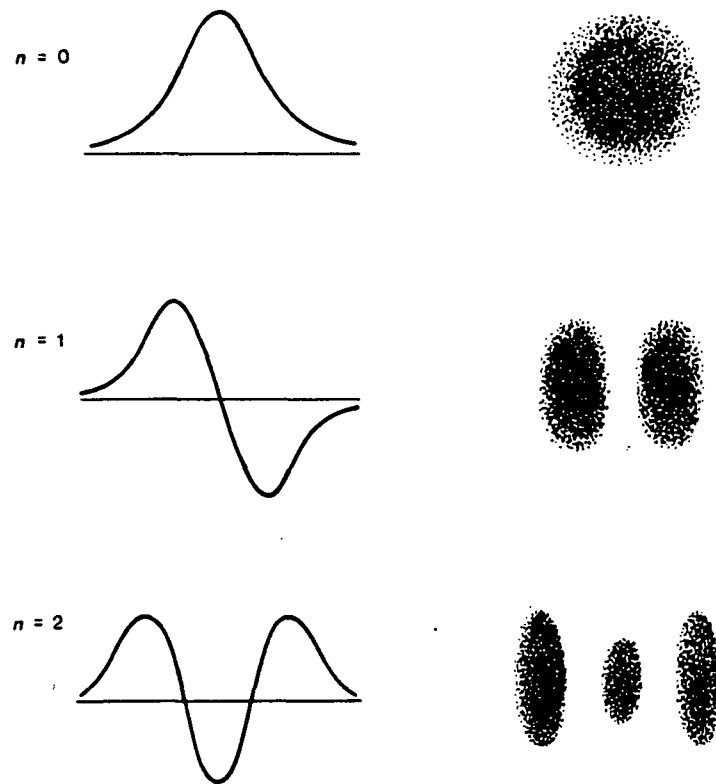
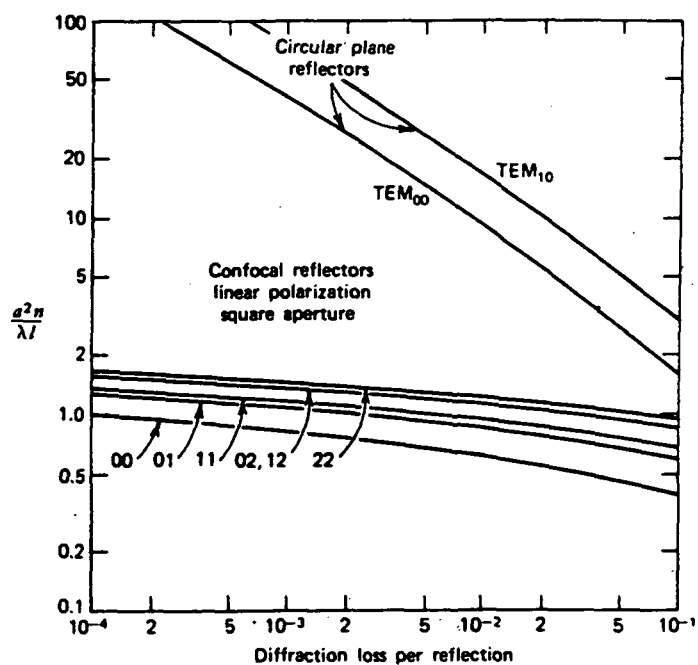


Figure 7-5 S-polarized light, showing the E-vector orientation, when reflecting off a conductor (top) and diffracting from a grating (bottom).



Hermite-gaussian transverse-mode patterns in a stable laser resonator.

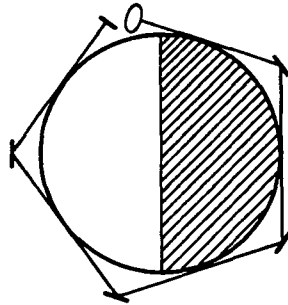
Figure 7-6 Field amplitude patterns of low order transverse cavity modes. [Siegman, 86]



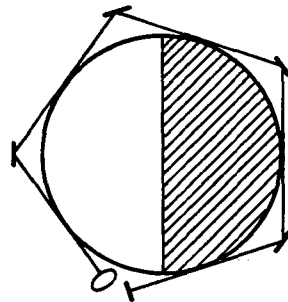
Diffraction losses for a plane-parallel and several low-order confocal resonators; a is the mirror radius and l is their spacing. The pairs of numbers under the arrows refer to the transverse-mode indices m, n .

Figure 7-7 Laser system Fresnel number versus mirror diffraction loss, parametrized by resonator type and transverse cavity mode. [Yariv, 75]

Bright sectors
at the 1α end.



Bright sectors
at the 1β end.



Bright sectors
at opposite ends.

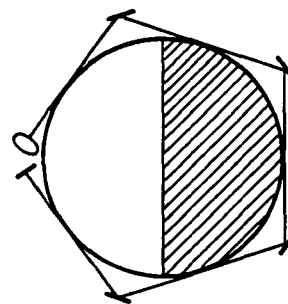
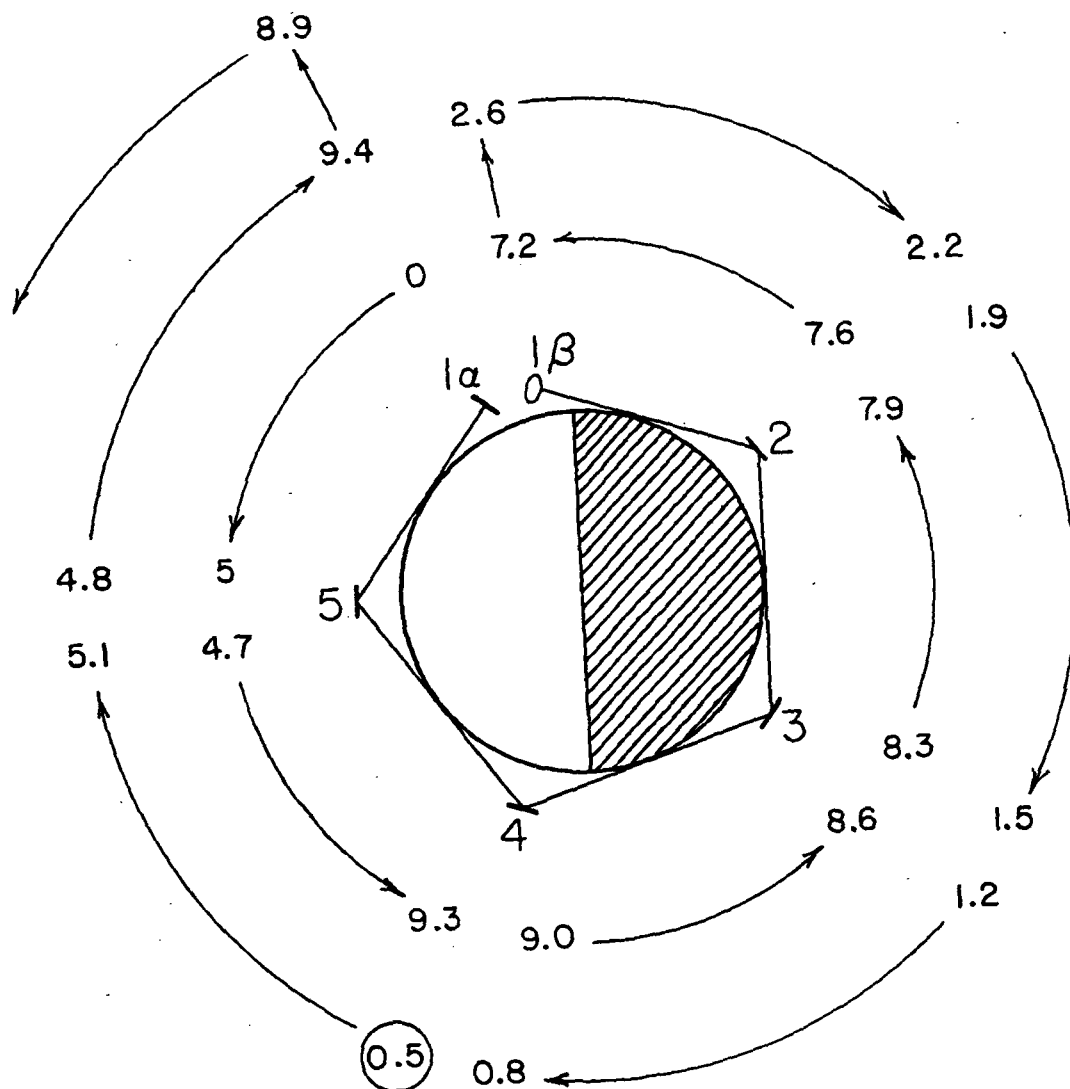


Figure 7-8 Resonator ring positions at various orbital times, showing critically different relationships between the ring gap and gain sectors.



Loss Source

diffraction

pointing

reflection (2,3,4,5)

(1α)

(1β)

grating inefficiency

output coupling

%

0.2 per sector

0.2 per sector

0.3 per surface

0.5

0.6

2

2

Figure 7-9 Cavity loss accounting.

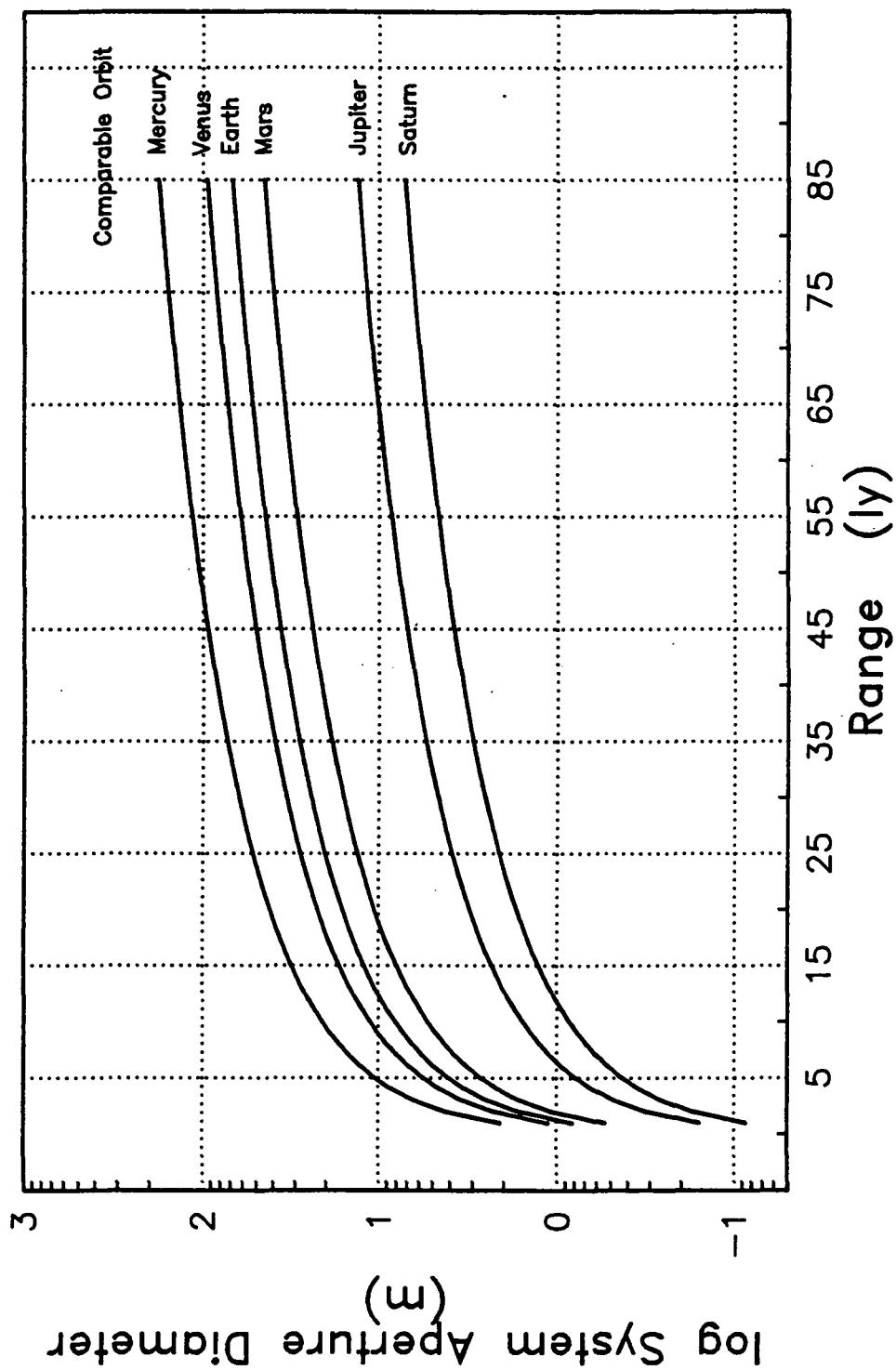


Figure 7-10 Final system aperture size required for diffraction spreading alone to cover various orbit sizes at interstellar distance.

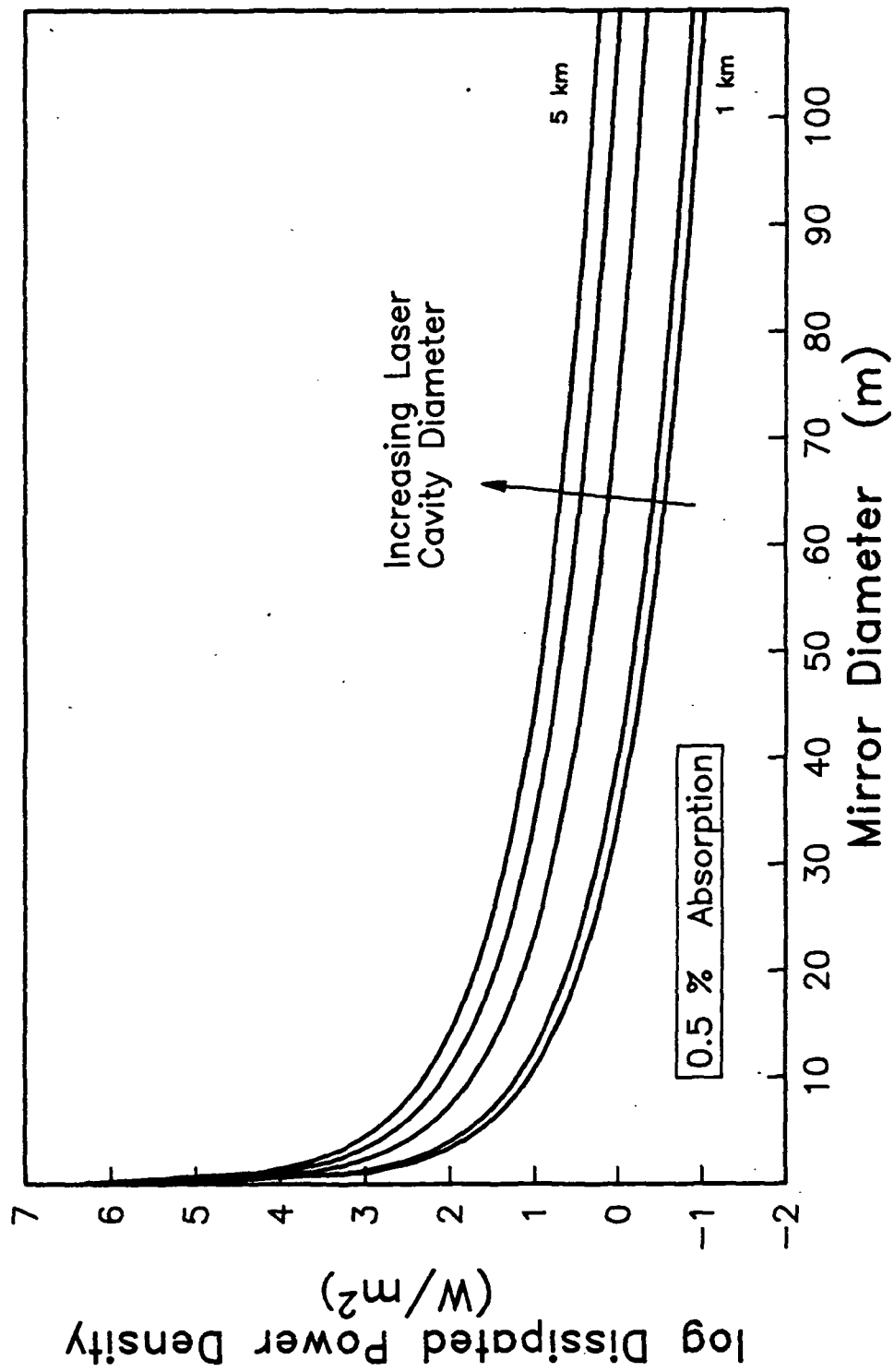


Figure 7-11 Heat load incurred by reflective optics in the path of the Venus laser output beam, assuming all non-reflected energy is absorbed.

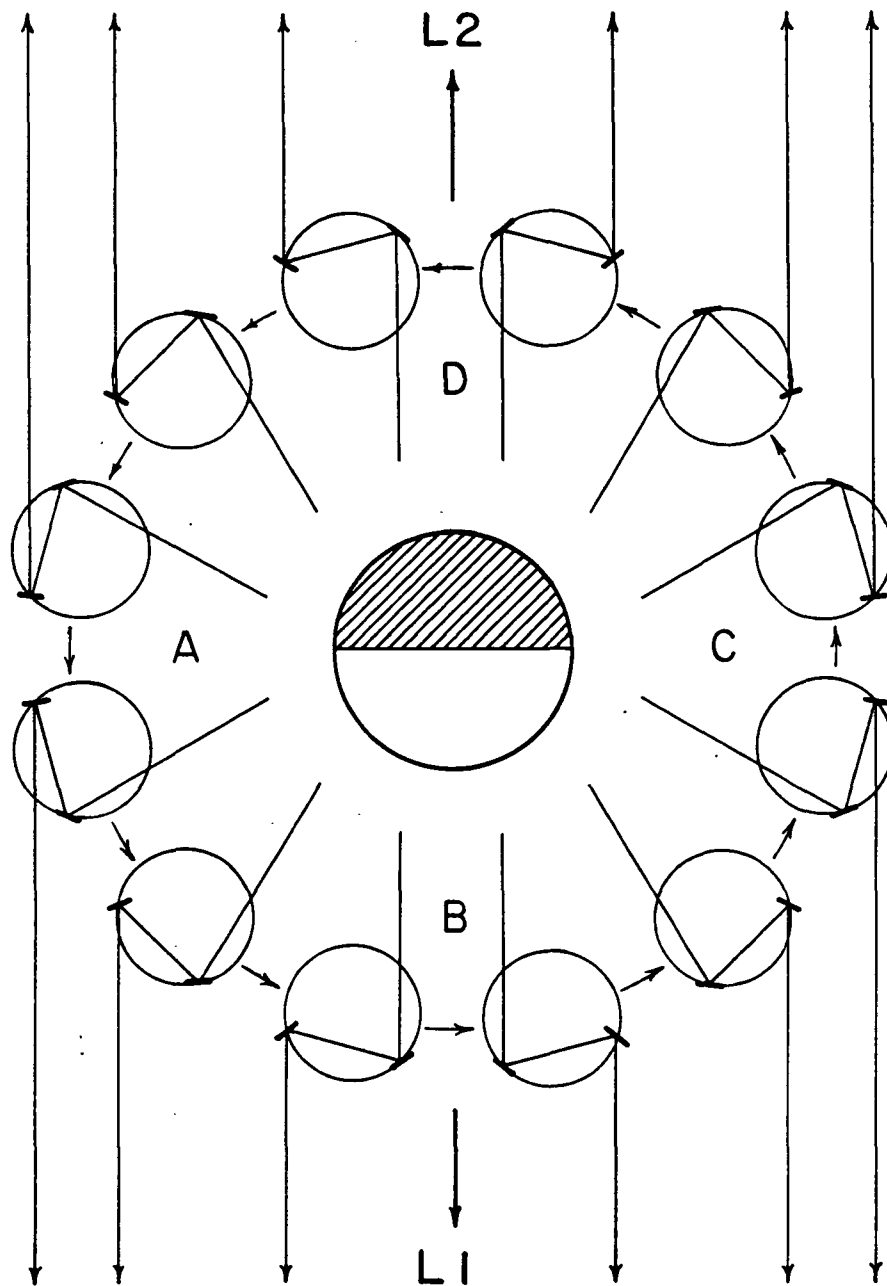


Figure 7-12 Tracking the libration point stations. The l_{ϵ} satellite rotates prograde 1 rev/orbit. At points A, B, C and D satellite l_{δ} switches (56 mrad) to l_{ϵ} 's opposite rim mirror. Simultaneously, at points A and C the rim mirrors switch the beam between L1 and L2; at B and D they pivot 45° to maintain contact with the respective modulator.

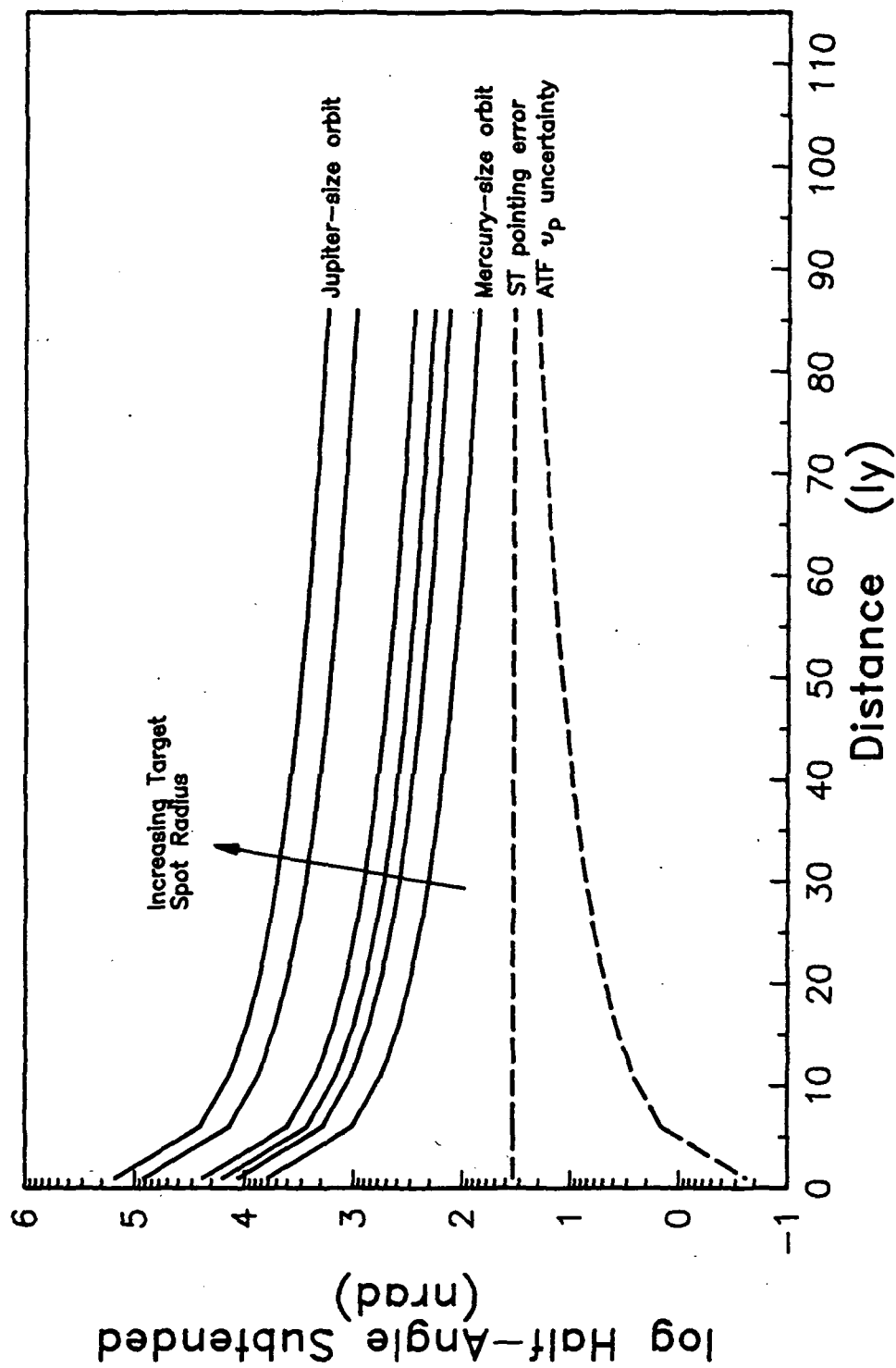


Figure 7-13 A comparison of the beam divergence required to cover distant orbits with reference angular uncertainties of pointing ability and star location.

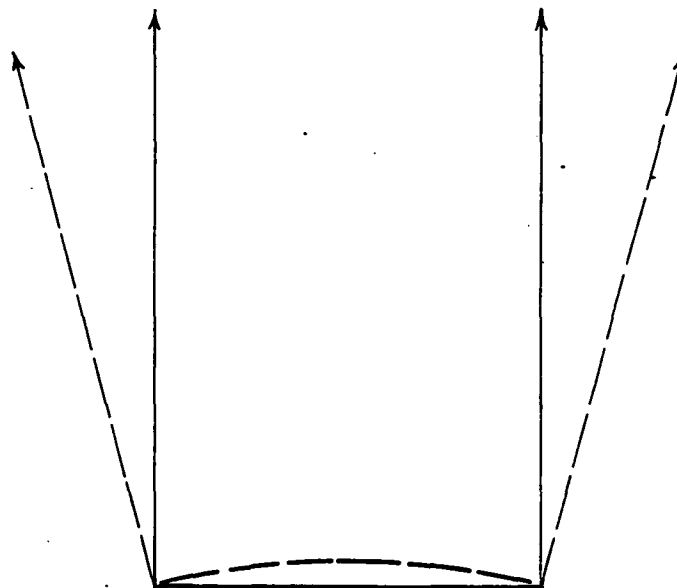
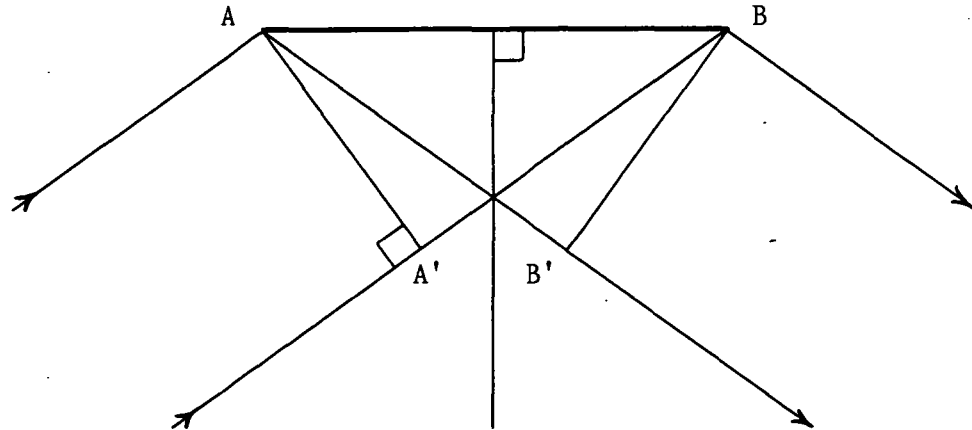


Figure 7-14 Schematic of an acousto-optic modulating mirror which employs a piezoelectric ceramic actuator stack to deflect a metal membrane, showing the divergence of surface normals and beam spreading which results.

Appendix A7-1 Proof of extreme ray congruence for polygonal resonator.

Single Plane Reflection



Since the angle of incidence must equal the angle of reflection, the entire geometry is reflectively congruent about the mirror normal. The incident plane wave front AA' remains plane as BB' after reflection because the extra distance $A'B$ traveled by the lower extreme ray equals the extra distance AB' traveled by the upper extreme ray.

Double Plane Reflection

The path congruence in this case can be proved by successive application of the above argument, or can be proved directly. Although the following proof is more cumbersome than the preceding one, it applies obviously to the double-reflector satellite geometry. The result is unchanged in any case: with plane mirrors, all parallel cavity rays travel identical distances, regardless of which is "innermost" or "outermost", or even if these reverse.

Referring to the following diagram, symmetry allows us to prove the case if the rays AOB and $A'O'B'$ are congruent.

Appendix A7-2 Focusing and shielding of micrometeoroid flux by Venus.

Gravitational focusing

Because Venus and Earth are in the same solar system neighborhood and are about the same size and mass, we obtain a rough estimate by adapting Earth data from [Griffin, 86]. Our orbital radius is:

$$\frac{\text{orbital radius}}{\text{venus radius}} = \frac{7641}{6052} = 1.26 \text{ planetary radii}$$

which yields a focusing factor of about 0.9 for interplanetary particles with average speed 20 km/s. We reduce this factor to account for Venus' smaller mass:

$$0.9 \frac{\mu_{\text{v}}}{\mu_{\oplus}} = 0.9 \left(\frac{3.257(10^5)}{3.986(10^5)} \right) \approx 0.7$$

We would therefore expect:

$$\frac{0.7}{0.57} = 1.3 \text{ times the deep space particle flux.}$$

Geometrical shielding

$$\sin\theta = \frac{\text{venus radius}}{\text{orbital radius}} = \frac{6052}{7641} = .7920$$

Therefore $\cos\theta = .6105$, and the body shielding factor ζ is:

$$\zeta = \frac{1 + \cos\theta}{2} = .805$$

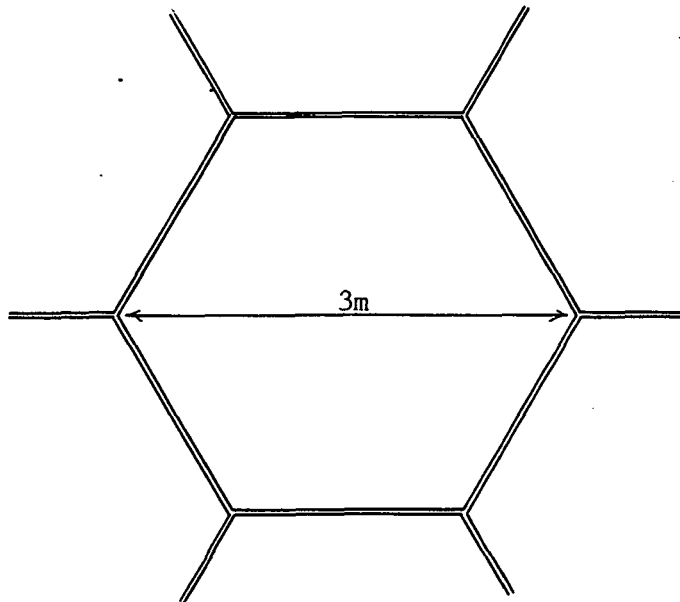
Ultimately then we expect 81 % of the focused flux, or about 1.05 times the deep space flux. The anisotropic compensation of shielding for focusing implies enhanced particle flux on the mechanical overbody and reduced flux on the optical underbelly of single-surface reflectors.

Appendix A7-3 Estimated mirror segment design.

Perkin-Elmer scaled up their 9.5" demonstrator to design a 36" mirror based on 21.1 kg/m^2 . They calculated a 14.1 kg mass, a total thickness of .052 m, a first resonance of 2000 Hz, and a self-weight deflection of .02 μm for this larger size. So confident were they in their process that they predicted halving the mirror weight "with minor modifications to the design" [Paquin et al, 84], even though their large mirror represents an area increase of:

$$\left(\frac{36}{9.5}\right)^2 = 14.4 \text{ over their demonstrator.}$$

We will scale up their process somewhat less than another order of magnitude, to hexagonal mirrors 3 m across in their long dimension.



This represents a further area increase of:

$$\frac{12\left(\frac{1}{2}(1.5 \cos 30^\circ)(1.5 \sin 30^\circ)\right)}{\pi((18)(.0254))^2} = 8.9$$

to an area of 5.85 m^2 , and a change in outline to allow tiling a plane.

Clearly a mass figure based on areal density cannot be used to arbitrarily large dimensions for a plate structure if it is to remain stiff. On the other hand, Perkin-Elmer were casual in their efforts to hone the honeycomb design. We expect an optimized design for space-based manufacturing, and permit relaxing their surface figure accuracy an order of magnitude to IR standards. To be conservative, instead of halving the areal density we will increase it 50 %, to 30 kg/m^2 . The basic mirror mass will therefore be taken as:

$$(5.85)30 = 175 \text{ kg}$$

which we will round up to 180 kg to allow for coating, and extra material at the attachment points on the rear.

Assume a 3-point back support, as does Perkin-Elmer. Such mirror supports are typically located at the radius of equilibrium R_E [Yoder, 86]:

$$R_E = \frac{\sqrt{2}}{2} R$$

which would be .32 m for the 36" version. For our hexagon, define an effective radius R_e as the radius of the circle of equal area:

$$R_e = \left(\frac{5.85}{\pi} \right)^{\frac{1}{2}} = 1.36 \text{ m}$$

so R_E for our mirror is .96 m.

Since we have allowed three times as much structural mass per unit area of mirror as Perkin-Elmer's state of the art, we can get a rough idea of where all that extra mass is going by estimating a thickness for the mirror sandwich. Treat the sandwich structure as a monolithic circular plate with overall material properties E (Young's modulus) and ν (inverse Poisson ratio). Roark [65] gives the maximum center

deflection y of such a plate of thickness t as:

$$y = - \frac{3W(m-1)(5m+1)r^2}{16\pi E_m^2 t^3}$$

where r is the radius, and the simply supported plate is uniformly loaded by total weight W . (Although our plate is not continuously supported, the expression we will derive would result even from a fixed-edge formulation, and so is not critically dependent on the displacement boundary conditions.) For two similar plates to experience the same maximum deflection, then:

$$\frac{Wr^2}{t^3} = \frac{W_0 r_0^2}{t_0^3}$$

from which:

$$t^3 = \frac{Wr^2}{W_0 r_0^2} t_0^3 \quad (\text{A7-3.1})$$

W is the mirror self-weight in this case, which in orbit becomes the body inertia force produced by acceleration, $W = ma$. For simplicity assume a constant acceleration requirement for both mirrors, so that:

$$\frac{W}{W_0} = \frac{m}{m_0}$$

Using R_E as the supported plate radius, and substituting mass and R_E values for Perkin-Elmer's 36" model and our 3 m hexagon, we get:

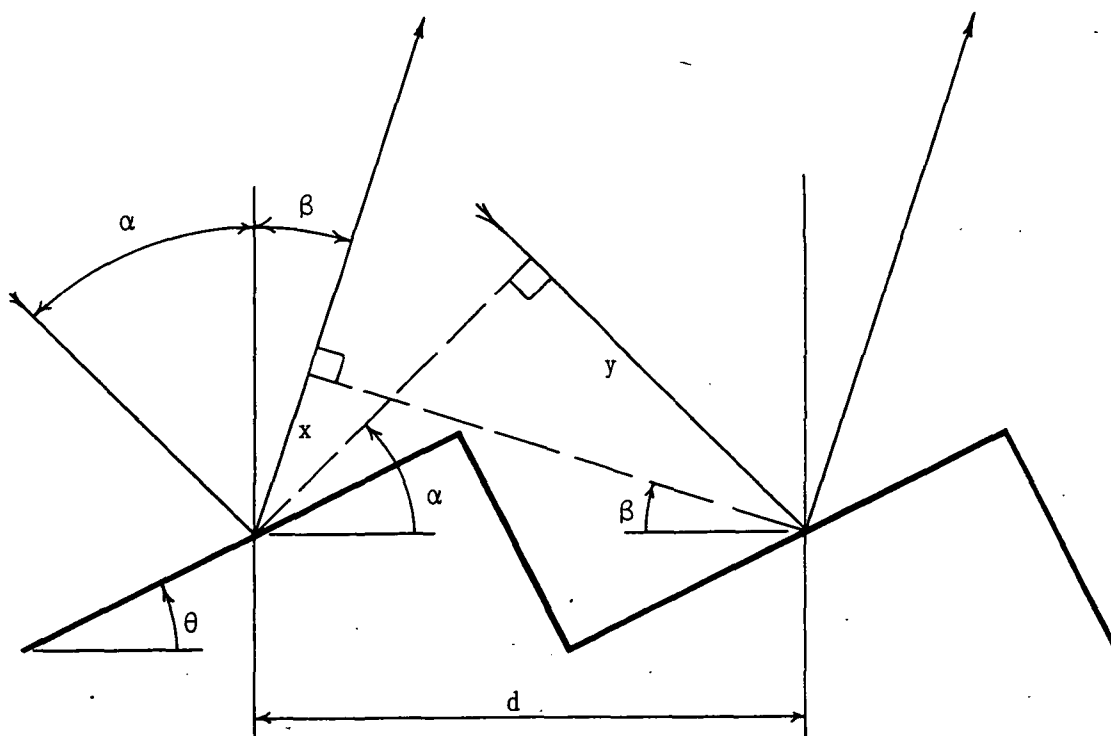
$$t = \left[\frac{180 (.96)^2}{14.1 (.32)^2} \right]^{1/3} (.052) = .25 \text{ m}$$

for ours, almost five times thicker than the Perkin-Elmer model.

Increasing the total thickness by five using only three times as much material indicates the amount of structural optimization required in designing the honeycomb core and facesheets for our larger case. The

first resonance should be about $2000 \left(\frac{.46}{1.5} \right) \approx 600 \text{ Hz}$.

Appendix A7-4 The grating equation.



The path length difference for rays on two adjacent rulings is simply:

$$\pm(x - y) = \pm(d \sin \alpha - d \sin \beta)$$

In general then, light diffracted from different grooves will be in phase when the path length difference equals an integral number of wavelengths, leading directly to the grating equation:

$$m\lambda = d(\sin \alpha \pm \sin \beta) \quad (\text{A7-4.1})$$

The integer m is called the diffraction order; see Figure 7-3 for its sign convention. $m = 0$ implies $\alpha = -\beta$, the degenerate case of specular reflection from a plane mirror.

Littrow configuration

If the grating is used such that the order $m = 1$ dominates, then $\alpha = \beta$ and the diffracted ray retraces the incident path. The grating equation reduces to:

$$m\lambda = 2d \sin\alpha \quad (A7-4.2)$$

Appendix A7-5 Proof of allowable grating orders for restricted groove spacing.

A grating designed for blazed Littrow use will have $\alpha = \beta \equiv \theta$ for the order $m = 1$. Equation A7-4.2 yields:

$$\sin\theta = \frac{\lambda}{2d} \quad (\text{A7-5.1})$$

If such a grating is indeed operated at its blaze angle, we can ask the question: as d is hypothetically varied, which general orders can the grating support? Substituting $\sin\theta$ from equation A7-5.1 into the general grating equation, A7-4.1, yields:

$$n\lambda = d \left(\frac{\lambda}{2d} + \sin\beta \right)$$

for integer n , which when rearranged specifies that:

$$\sin\beta = \frac{(n - \frac{1}{2})\lambda}{d} \quad (\text{A7-5.2})$$

Two cases are possible, corresponding to whether the diffracted ray is on the same side of the grating normal as the incident ray, or on the opposite side.

Case 1: for $\beta \geq 0$, $0 \leq \sin\beta \leq 1$

Inserting equation A7-5.2 into the first half of this inequality leads directly to $\frac{1}{2} \leq n$, which can be rewritten $1 \leq n$ since n must be an integer. Working with the second half of the inequality yields:

$$n \leq \frac{d}{\lambda} + \frac{1}{2} \quad \text{or simply} \quad d \geq \frac{\lambda}{2}$$

since n is bounded from below by 1.

If we let n be any integer > 1 , the second half of the inequality yields:

$$2 \leq \frac{d}{\lambda} + \frac{1}{2} \quad \text{or} \quad d \geq \frac{3\lambda}{2}$$

Thus the strict inequality $d < \frac{3\lambda}{2}$ implies $n < 2$, and we conclude that for rays diffracted on the incidence side of the grating normal, only the order $n = 1$ is permitted if:

$$\frac{\lambda}{2} \leq d < \frac{3\lambda}{2} \quad (\text{A7-5.3})$$

Case 2: for $\beta \leq 0$, $-1 \leq \sin \leq 0$

Substituting equation A7-5.2 into the second half of this inequality leads directly to $n \leq \frac{1}{2}$, rewritten as $n \leq 0$ since n is an integer. The first half of the inequality yields:

$$-\frac{d}{\lambda} + \frac{1}{2} \leq n \quad \text{or simply} \quad d \geq \frac{\lambda}{2}$$

since n is bounded from above by 0.

Letting n be any integer < 0 , the first half of the inequality yields:

$$-\frac{d}{\lambda} + \frac{1}{2} \leq -1 \quad \text{or} \quad d \geq \frac{3\lambda}{2}$$

Thus the strict inequality $d < \frac{3\lambda}{2}$ implies $n > -1$, and we conclude that for rays diffracted on the opposite side of the grating normal from the incident ray, only the order $n = 0$ is permitted if the groove spacing d conforms to equation A7-5.3.

Appendix A7-6 CO₂ rotational distribution.

The population of CO₂ molecules in any given rotational state J as a fraction of the total population is given by [Herzberg, 66]:

$$n_J = \frac{2(2J + 1) \exp\left(-\frac{BhcJ(J + 1)}{kT}\right)}{\sum_J n_J}$$

where the J are restricted to even integers because CO₂ is a linear symmetrical triatomic molecule, B ≡ 1st coefficient of molecular energy series expansion, a rotational constant unique to the molecule (.390 cm⁻¹ for CO₂), h ≡ Planck constant (6.626(10⁻³⁴) Js), c ≡ vacuum speed of light (2.998(10¹⁰) cm/s), k ≡ Stefan-Boltzmann constant (1.381(10⁻²³) J/K), and T ≡ gas kinetic temperature.

The rotational distribution varies only with molecular species and temperature. At the gain altitude of 130 km on Venus, the temperature may be taken as 195 K (Appendix A6-7).

The FORTRAN program LSRLIN.FTN which follows uses these values to calculate the CO₂ rotational distribution in our laser medium.

C-5

```

ftn,1,s
$files 0,1
  program lsrlin
  real n(61),f(61)
  open(99,file = 'rotpop.dat')

  c = -0.0029

  do 2 i = 1,61,2
    j = i-1
    f(i) = 2*(2*j+1)*exp(c*j*(j+1))
    t = t + f(i)
2  continue

  do 3 i = 1,61
    j = i-1
    n(i) = 100.*f(i)/t
    write(99,*) j,n(i),t,f(i)
3  continue

  close(99)
  stop
  end

```

Appendix A7-7 Diffraction grating design.

For a blazed grating in the Littrow configuration, the groove spacing d is found for the order $m = 1$ from equation A7-4.2:

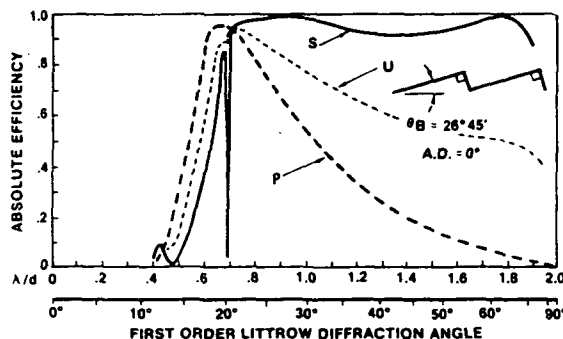
$$d = \frac{\lambda}{2\sin\theta}$$

which when substituted into the groove spacing interval specified by equation A7-5.3 yields $1 \geq \sin\theta > 1/3$, or:

$$90^\circ \geq \theta > 19.47^\circ$$

if only the $m = 1$ and $m = 0$ orders are to be permitted. For this reason, high blaze-angle gratings ($22^\circ < \theta < 38^\circ$) are widely used for laser tuning.

The graph, taken from [OISPD, 84], shows the absolute efficiency (the fraction of light diffracted into the order $m = 1$ in this case) for a grating with $\theta = 26.75^\circ$, calculated by solving Maxwell's equations with the boundary

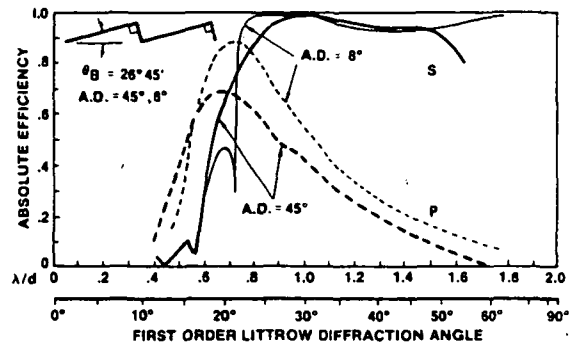


conditions of a perfect grating of infinite conductivity (perfectly reflecting). Since only $m = 1$ and $m = 0$ are allowed, the fraction of energy not diffracting back to the source ($m = 1$) can be presumed to diffract at the degenerate specular angle ($m = 0$). Actual performance (relative efficiencies) for these orders would simply be their absolute efficiencies multiplied by actual reflectivity (< 1). High blaze-angle gratings feature a broad peak of near 100 % efficiency for s-polarized light, providing a generous and predictable tuning range for lasers.

This graph shows what happens to the grating's absolute efficiency if it is detuned from the strict Littrow configuration by small (8°) and large (45°) angles.

S-polarized efficiency remains extremely high for the design wavelength, which for us means

that the slight angular detuning required for orbital point-ahead will not detract significantly from diffraction efficiency.



Choosing $\theta = 26.75^\circ$ for our grating, then, implies $\lambda/d \approx .90$, so that for $\lambda = 10.513 \mu\text{m}$, $d = 11.68 \mu\text{m}$, a completely ordinary groove specification, even for conventional ruling techniques.

Appendix A7-8 Line tuning resolution of design grating.

Assume the grating of Appendix A7-7 is oriented with a nominal incidence angle $\alpha = 26.75^\circ$. The nearest CO_2 spectral lines to the P(12), the P(10) and the P(14), both differ by $.019 \mu\text{m}$ in wavelength from it. Assume the P(14), with wavelength $\lambda = 10.532 \mu\text{m}$. Then for the order $m = 1$, equation A7-4.1 yields:

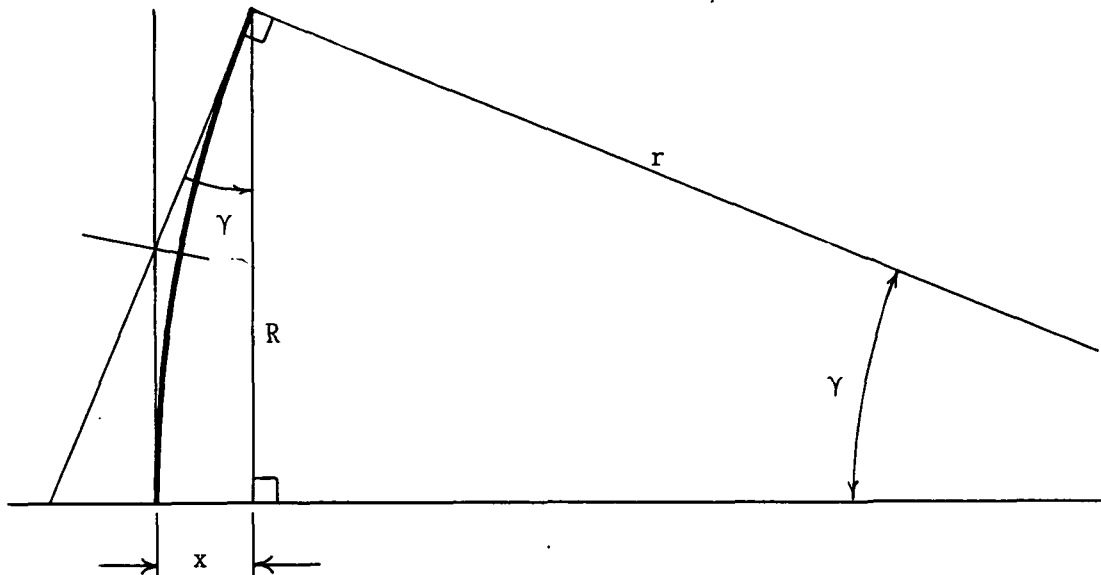
$$\begin{aligned}\beta &= \sin^{-1} \left[\frac{\lambda}{d} - \sin \alpha \right] \\ &= \sin^{-1} \left[\frac{10.532}{11.68} - \sin 26.75^\circ \right] = 26.85^\circ\end{aligned}$$

Any emergent energy on this adjacent CO_2 line will therefore miss the Station 2 reflector by:

$$8983 \tan(26.85^\circ - 26.75^\circ) = 16.3 \text{ km}$$

exiting the oscillator immediately and remaining unamplified.

Appendix A7-9 Reference figure deviation for shallow confocal mirror.



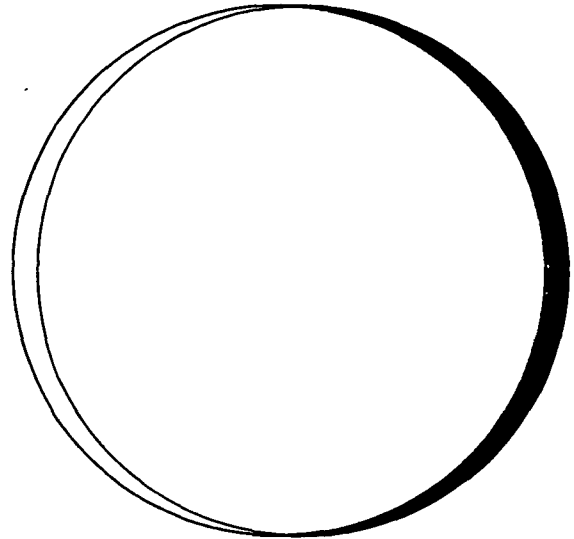
$$\frac{R}{r} = \sin \gamma \approx \frac{2x}{R}$$

So
$$x \approx \frac{R^2}{2r} = \frac{500^2}{(2)8983(10^3)} \approx .014 \text{ m}$$

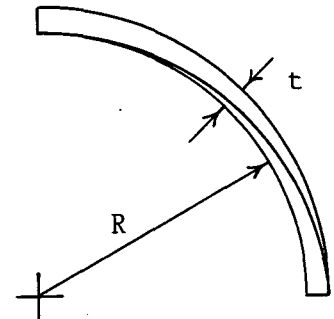
Thus 1 km diameter mirrors confocal for a separation of 8983 km would have a center-to-edge figure deviation, from plane of 1.4 cm.

Appendix A7-10 Pointing accuracy requirement for cavity satellites.

The effect of a slight intracavity pointing error would be the loss of a thin crescent of the beam cross section. Limiting that loss to 0.2 % per pass means limiting the crescent area to .002 times the nominal beam area.



The crescent area can be approximated by flipping the bottom half up around the top half to form a quarter ring of constant thickness t . Setting this area equal to the budgeted loss:



$$\frac{2\pi R t}{4} \equiv (.002)\pi R^2$$

and rearranging yields:

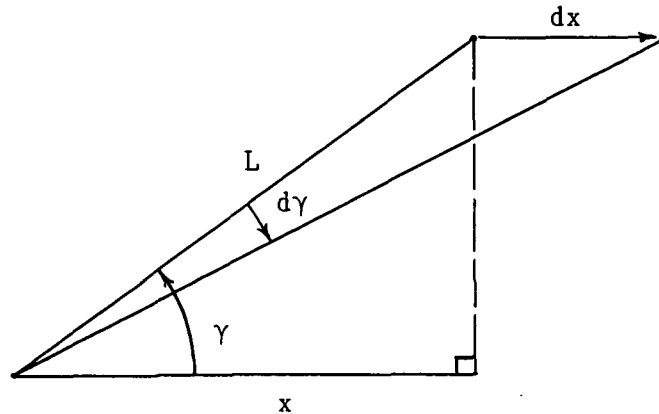
$$t = (.004)R = 2 \text{ m}$$

which, as the maximum crescent thickness, represents the lateral pointing offset resulting in the specified loss.

Over the separation distance, this lateral error implies an angular error of:

$$\frac{2}{8983(10^3)} = 0.22 \text{ } \mu\text{rad}$$

Appendix A7-11 Resonator satellite pointing biases.



$$\cos \gamma = \frac{x}{L}$$

Implicit differentiation yields:

$$-\sin \gamma d\gamma = \frac{dx}{L}$$

The light travel delay between vertex stations is .03 s, during which time the target station moves at 6.52 km/s along its orbit. That distance dx is here represented as a straight line (accurate to $2(10^{-5})$). The angle $\gamma = 36^\circ$, the complement to the nominal mirror incidence angle. Substituting these values gives:

$$|d\gamma| = \frac{dx}{L \sin \gamma} = \frac{(.03)(6.52)}{8983 \sin 36^\circ} = 37 \text{ } \mu\text{rad}$$

which is the pointing offset bias required in general for the light leaving one station to arrive at the next station's updated location. When effected by mirror tilt, this point-ahead bias works for both ray directions, allowing cavity continuity despite the ring's orbital motion.

Because the constellation of Station 1 is distributed geometrically about the pentagonal vertex it occupies, the lines of sight between Stations 1 α and 5, and Stations 1 β and 2, depart 300 m laterally over the course of their length from the nominal pentagon. This amounts to a required pointing adjustment for stations 2 and 5 of 33 μ rad. For Station 2, this value will subtract from the general 37 μ rad bias, and for Station 5, it will add to that general bias.

The actual mirror angle biases will for all satellites be half of the pitch biases just evaluated, since the mirror normal must bisect the opening angle between incident and reflected rays. Thus the reference values for satellite mirror angle biases will be:

Stations 1 β and 2	-2 μ rad
Stations 3 and 4	-18.5 μ rad
Stations 5 and 1 α	-35 μ rad

Appendix A7-12 Planetary laser oscillation linewidth.

Natural Doppler-Broadened Width

The full-width half-maximum (FWHM) width of a Doppler-broadened laser gain profile is given as [Siegman, 86]:

$$\begin{aligned}\Delta\nu_d &= \left(\frac{(8 \ln 2) kT}{Mc^2} \right)^{\frac{1}{2}} \nu_0 & (A7-12.1) \\ &= \left[\frac{(8 \ln 2) 1.38(10^{-23}) 195 (10^3)}{\left[\frac{44}{6.02(10^{23})} \right] (3(10^8))^2} \right]^{\frac{1}{2}} 28.5(10^{12}) \\ &= 43 \text{ MHz}\end{aligned}$$

for CO_2 (atomic mass $M = 44$ g/mole) in the Venusian mesosphere.

Ultimate Limit to Monochromaticity

The quantum-limited spectral purity is given by [Svelto, 82]:

$$\Delta\nu_{\text{osc}} = 4\hbar \frac{\nu_{\text{osc}}}{P} (\Delta\omega_c)^2 \quad (A7-12.2)$$

where $\nu_{\text{osc}} \equiv$ the center frequency ($28.516(10^{12})$ Hz), $P \equiv$ the output power (180 kW), and $\Delta\omega_c \equiv$ the cavity mode width, calculated as the inverse of the cavity photon decay time τ_c :

$$\Delta\omega_c \approx \frac{\gamma c_0}{L} \quad (A7-12.3)$$

where $\gamma \equiv$ total system losses (taken as equal to the single-circuit gain of 10 % for steady-state lasing), $c_0 \equiv$ the vacuum speed of light, and $L \equiv$ the single-circuit cavity length.

Substituting equation A7-12.3 into A7-12.2, we have:

$$\Delta\nu_{\text{osc}} = \left[\frac{(4)1.05(10^{-34})28.516(10^{12})}{1.8(10^5)} \right] \left[\frac{(.1)3(10^8)}{45(10^6)} \right]^2$$

$$\approx 3(10^{-26}) \text{ Hz}$$

Cavity Length Control

To oscillate, the laser field must repeat itself (be in phase with itself). Thus the coherence length must be at least twice the ring perimeter (the longest possible path goes from one end all the way through the cavity and then back again), or $89.8(10^6)$ m. Therefore the coherence time τ must be at least:

$$\tau = \frac{89.8(10^6)}{3(10^8)} = 0.3 \text{ s}$$

so the laser oscillation linewidth cannot exceed:

$$\Delta\nu = \frac{1}{\tau} = 3.3 \text{ Hz}$$

To see directly how oscillation frequency changes relate to cavity length changes, we note that the oscillation condition specifies that the cavity length contain an integral number of half-waves (that way the light's electric field vanishes at the cavity ends, allowing the field to repeat itself):

$$L = \frac{n\lambda}{2} = \frac{nc}{2\nu} \quad (\text{A7-12.4})$$

where the integer n is called the cavity longitudinal mode number. We are interested in how the cavity changes with time, which can be

represented through differentiation as:

$$\frac{dL}{dt} = \left(\frac{c}{2\nu} \right) \frac{\partial n}{\partial t} + \left(- \frac{nc}{2\nu^2} \right) \frac{\partial \nu}{\partial t} \quad (\text{A7-12.5})$$

Assume first that the mode number n is time-invariant, so the first term of equation A7-12.5 vanishes. Substituting n in terms of L (from equation A7-12.4) into the coefficient of the second term and simplifying yields:

$$\left| \frac{dL}{dt} \right| = \frac{L}{\nu} \frac{\partial \nu}{\partial t} \quad (\text{A7-12.6})$$

If we require the frequency to be stable within 3.3 Hz over the time necessary for the field to repeat itself (the double-perimeter light delay of .300 s), then we find that the total cavity length must not change by more than:

$$\frac{(5)8983(10^3)}{28.516(10^{12})} \left(\frac{3.3}{.3} \right) = 17.3 \text{ } \mu\text{m/s}$$

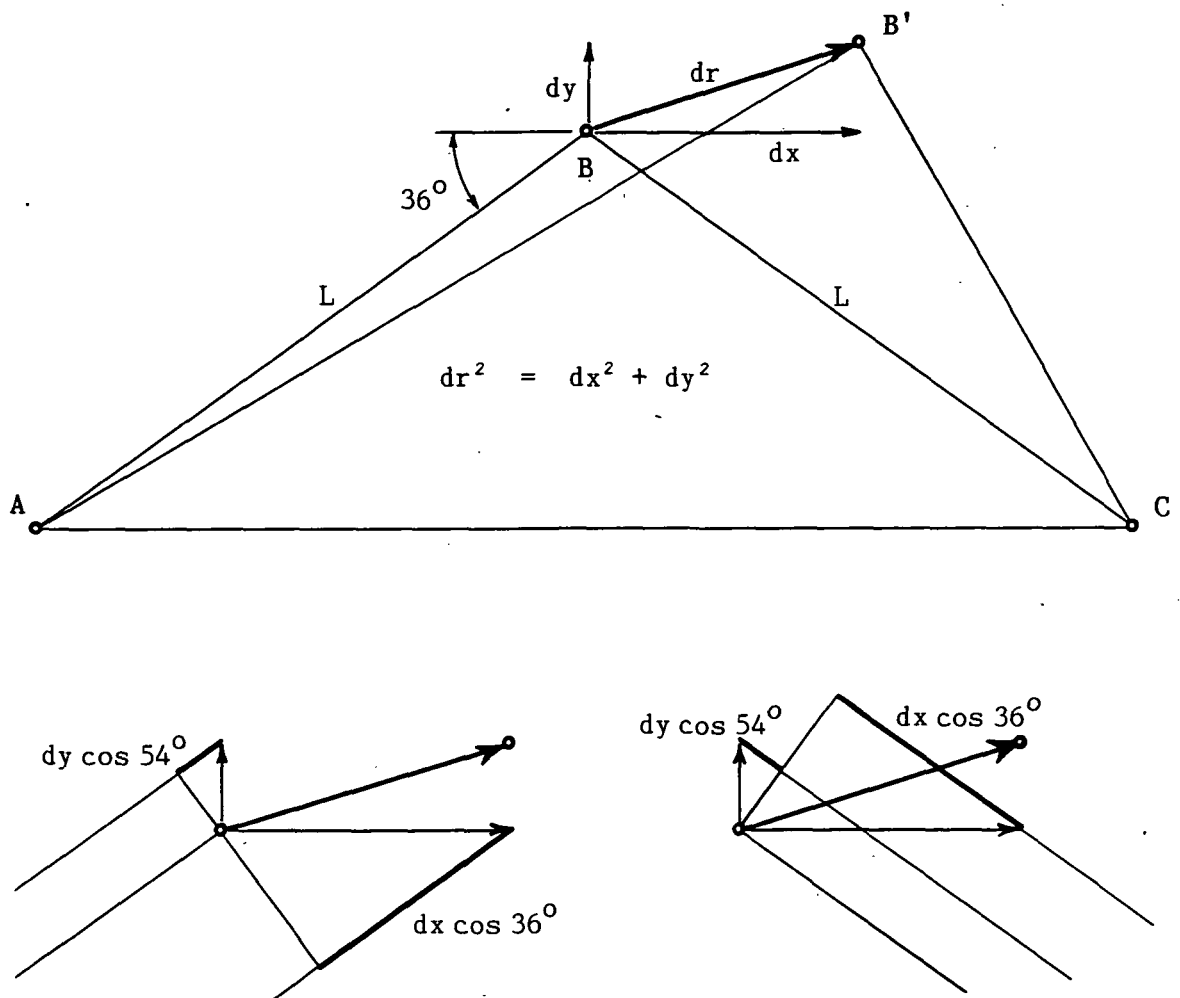
if laser oscillation is to occur.

Bearing in mind that the resonator satellites move at 6.52 km/s, and that we expect kilometer-scale excursions from their nominal positions due to gravity field irregularities, we might well suspect that a planetary laser would be inoperable, given the coherence length constraint just developed. However, the mode number n is not time-invariant. Imagine that the cavity length increases. If n remains constant, then the half-waves which fit into L an even n times must increase also, so their frequency ν decreases, as the negative sign in equation A7-12.5 shows. But as soon as the cavity length increases by $\lambda/2$ (5.257 μm in our case), n will jump to $(n + 1)$ since the cavity can now accommodate an "extra" half-wave in its length L . Thus the laser will not notice cavity length changes as long as they

result in an integral n for the oscillating wavelength.

Hence the planetary laser's mirror control system must insure that, no matter how the satellites themselves move, the optical path within the cavity only sees those motions discretized into jumps of $\lambda/2$. The positional accuracy of those steps must of course be $\lambda/20$, or $0.5 \mu\text{m}$ in our case, and the margin for drift error is the $17.3 \mu\text{m/s}$ rate just calculated. Should the cavity length change, uncompensated by mode jumps and continuous fine adjustment, at a rate faster than $17.3 \mu\text{m/s}$, lasing will cease because the intracavity beam will lose its temporal coherence.

We can translate this optical requirement into a performance specification for the satellite hardware.



The diagram represents one of the basic vertex stations (2, 3, 4, 5) as deviating from its nominal position B along the vector dr to B' . A coordinate system is set up such that the x-direction lies normal to, and the y-direction parallel to, that station's orbit radius. The components dx and dy of dr each contribute to changing the path length L as the beam both arrives and departs. Those changes are evaluated in the sub-diagrams, and the total path length change can be calculated by subtracting the original path length $2L$ from the sum of segments $\overline{AB'}$ and $\overline{B'C}$:

$$\begin{aligned}\Delta L &= (L + dy \cos 54^\circ + dx \cos 36^\circ) \\ &\quad + (L + dy \cos 54^\circ - dx \cos 36^\circ) - 2L \\ &= 2 dy \cos 54^\circ\end{aligned}\tag{A7-12.7}$$

For these four vertex stations, then, only motion normal to the orbit tangent affects the cavity length. However, motions parallel to the orbit tangent by the end stations (1α and 1β) do affect cavity length, because when the arriving and departing beams essentially coincide, their dx contributions add instead of cancel. The greatest change in path length will occur for these satellites when dr is parallel to the path, so for these:

$$\Delta L = 2 dr\tag{A7-12.8}$$

In all cases, we have only mirror actuation available to effect cavity length microcontrol, and it is the accuracy of this actuation that we seek to specify. An optimized system will result if all actuators yield identical performance. Call motion in the direction the actuator works dw . Then for Stations 2, 3, 4 and 5, $dw = dy$ because the mirror normals parallel the orbit radius. For Station 1α , $dw = dr$ because the mirror normals parallel the beam path. For Station 1β ,

the motion dw only contributes a component $dw \cos 26.75^\circ$ in the direction of dr because the grating mirrors are tipped at their blaze angle. Now the contributions to ΔL from all six resonator satellites cannot exceed $\lambda/20$ for resolution, so we have:

$$L = \frac{\lambda}{20} = 2dw + 2dw \cos 26.75^\circ + 4(2dw \cos 54^\circ)$$

from which:

$$dw = \frac{\lambda}{20 (2 + 2 \cos 26.75^\circ + 8 \cos 54^\circ)} = .0619 \mu\text{m}$$

is the required actuator position resolution. Similarly:

$$\frac{dw}{dt} = \frac{17.3}{(2 + 2 \cos 26.75^\circ + 8 \cos 54^\circ)} = 2.04 \mu\text{m/s}$$

is the maximum allowable uncompensated drift rate for mirror control.

As the resonator satellites move under the influence of perturbing forces then, a mirror control system which can actively "mode hop" while maintaining the accuracy specified above will satisfy the oscillation condition and allow lasing. The emergent beam will be extremely spectrally pure (3.3 Hz) about a base frequency which may wander by up to:

$$\frac{17.3(10^{-6})}{45(10^6)} 28.516(10^{12}) = 11 \text{ Hz/s}$$

A receiver with many adjacent narrowband channels could easily track such excursions, while allowing extremely high signal-to-noise ratios (Chapter 3). Thus the planetary laser's enormous length, which makes it difficult to operate, also ensures that it must be a tremendously effective tool for long-distance communication.

Appendix A7-13 Beam power.

Assuming steady-state power extraction equal to the solar pumping rate, the laser photon flux F can be taken as:

$$F \approx \beta V = \pi \beta L' R_c^2$$

where $\beta \equiv$ the volume emission rate of the inverted Venusian mesosphere, $R_c \equiv$ the laser cavity radius, and $L' \equiv$ the effective cavity length. The interaction length in any bright ring sector is about 400 km, so we introduce L' as a simple way of inflating this dimension commensurate with the increase in minimum system available gain afforded by the multi-sector ring. Figure 5-7 showed the pentagonal ring minimum gain to exceed the peak single pass value by 50 %, so we will set L' at 150 % the single-pass interaction length, or 600 km. Then:

$$F \approx \pi 2(10^{13}) 6(10^5) 500^2 = 9.4(10^{24}) \frac{\text{ph}}{\text{s}}$$

One photon of 10.5 μm light has energy:

$$h\nu = 6.626(10^{-34}) 28.516(10^{12}) = 1.890(10^{-20}) \text{ J}$$

so the output beam power is:

$$Fh\nu = 180 \text{ kW}$$

Assuming the output beam to represent a 2 % coupling loss, the circulating cavity power is:

$$(.18) \frac{98}{2} = 8.7 \text{ MW}$$

That large power is spread over the large beam cross section, however, so the intracavity power density is:

$$\frac{8.7(10^6)}{\pi(500)^2} = 11 \frac{\text{W}}{\text{m}^2}$$

which is less than 1 % of the 1.4 kW/m^2 intensity of sunlight at 1 AU.

Appendix A7-14 Program APERTURE.FTN for calculating system aperture required for diffraction spreading to cover distant planetary orbits.

The program presumes that only the central Airy spot is of interest for interstellar targeting. The radius of this spot at a distance L for light of wavelength λ leaving an aperture of diameter D is:

$$R_s = \frac{(1.22)\lambda L}{D}$$

```
ftn,1,s
$files 1,1
  program aperture

c system aperture required for diffraction limited
c coverage of distant planetary orbits

  real rs(6),l(85),d(6,85)
  open(99,file = 'aperture.dat')

c distant orbit radii

  rs(1) = 57.9e9
  rs(2) = 108.2e9
  rs(3) = 149.6e9
  rs(4) = 227.9e9
  rs(5) = 778.3e9
  rs(6) = 1427.e9
  do 2 i = 1,6

c vary link distance and calculate aperture diameter

    l(1) = 1.
    do 3 j = 1,85
      d(i,j) = 1.28e-5 * l(j) * 9.47e15 / rs(i)
3      l(j+1) = l(j) + 1.
2      continue

    do 4 k = 1,85
4      write(99,*) l(k), (d(i,k),i=1,6)

  close(99)
  stop
end
```

Appendix A7-15 Program PWRDEN.FTN for calculating the non-reflected power density on system mirrors of various size inserted in the output beam, parametrized according to laser cavity diameter (available power).

```

ftn,1,s
$files 2,2
    program pwrden

c power density at aperture mirror for diffraction limited
c performance

    real da(175),ga(175),dc(5),p(5),pd(5,175),gpd(5,175)
    open(99,file = 'pwrdenr.dat')
    open(98,file = 'pwrden.dat')
    pi = 4*atan(1.)

    dc(1) = 1.e3
    do 2 i = 1,5
        p(i) = .178 * dc(i)**2
        da(1) = .1
        do 3 j = 1,175
            pd(i,j) = .0064 * p(i) / da(j)**2
            ga(j) = log10(da(j))
            gpd(i,j) = log10(pd(i,j))
3        da(j+1) = da(j) + 1.
2    dc(i+1) = dc(i) + 1.e3

    do 4 m = 1,175
        write(98,*) da(m), (pd(i,m), i=1,5)
4    write(99,*) ga(m), (gpd(i,m), i=1,5)

    close(98)
    close(99)
    stop
end

```

Appendix A7-16 Beam pointing between the resonator and modulators.

Being slightly farther from Venus than L1, L2 represents the worst case. Were the intermediate beam simply a 10 m diameter plane wave, it would spread to a spot diameter of:

$$D_s = \frac{2(1.22)\lambda L}{D_A} = \frac{2(1.22)10.5(10^{-6})1.014(10^9)}{10}$$

$$= 2600 \text{ m}$$

between the 1ε switch and L2. An optic at L2 of projected diameter 10 m would therefore intercept only:

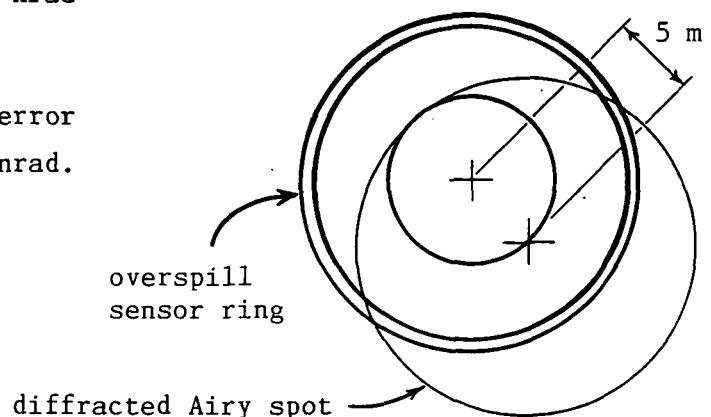
$$(.84) \frac{5^2}{1300^2} = 1.2(10^{-5})$$

of the intermediate beam's energy, since the central Airy spot we are calling the beam contains itself only 84 % of the total energy.

Specifying an Airy spot size twice the intercepted diameter, or 20 m, allows capturing 25 % of the spot energy (21 % of the total beam energy) and requires a lateral pointing error smaller than 5 m. The angular pointing accuracy required of the beam as it leaves 1ε is:

$$\frac{5}{1.014(10^9)} \approx 5 \text{ nrad}$$

The Hubble Space Telescope error specification is about 35 nrad.



Appendix A7-17 Program ANGERR.FTN to calculate pointing errors due to uncertainties in spacecraft stability, target location and target size.

```
ftn,1,s
$files 1,1
      program angerr

c compares optimal half beam spread with system errors

      real ltyr(18),rs(6),hbs(18,6),atf(18)
      open(99,file = 'angerr.dat')

      st   = 34.
      rs(1) = 57.9e9
      rs(2) = 108.2e9
      rs(3) = 149.6e9
      rs(4) = 227.9e9
      rs(5) = 778.3e9
      rs(6) = 1427.e9

      ltyr(1) = 1.
      do 2 i =1,18
          atf(i) = .24 * ltyr(i)
          do 3 j = 1,6
3          hbs(i,j) = (rs(j)/(ltyr(i) * 9.47e15)) * 1.e9
2      ltyr(i+1) = ltyr(i) + 5.

      do 4 k = 1,18
4      write(99,*) ltyr(k),st,atf(k),(hbs(k,m),m=1,6)

      close(99)
      stop
      end
```

Appendix A7-18 Required mirror deflection to effect proper beam spread..

Assuming a target the size of Earth's orbit at a distance of 10 ly, the necessary beam half-angle is:

$$\frac{149.6(10^9)}{(10)9.47(10^{15})} = 1.6 \text{ } \mu\text{rad}$$

Using the geometry of Appendix A7-9, the center-to-edge mirror deflection x is:

$$x \approx \frac{\gamma R}{2} = \frac{1.6(10^{-6})7.5}{2} \approx 6 \text{ } \mu\text{m}$$

for a 15 m mirror to spread the beam sufficiently. Such translation and its attendant accuracy requirement are easily attainable with a variety of piezoelectric or electromagnetic pushers.

References

C W Allen- (73) [op cit Chapter 1].

American Institute of Physics Handbook 3rd ed
(McGraw-Hill, 1972).

William P Barnes Jr- "Basic Properties of Metal Optics"
Optical Engineering Vol 16 No 4 (Jul-Aug 1977).

Bausch & Lomb Diffraction Grating Research Laboratory-
Diffraction Grating Handbook (Bausch & Lomb, 1970).

Edward H Bock- Lunar Resources Utilization for Space
Construction Final Report in 3 volumes NASA CR-173024
(30 Apr 1979).

William J Boyer (ed)- Large Space Antenna Systems Technology
1984 NASA Conference Publication 2368 (1985).

Arthur C Clarke- Profiles of the Future (Warner, 1984).

M G Cohen & E I Gordon- "Electro-Optic Gratings for Light
Beam Modulation and Deflection" Applied Physics Letters
Vol 5 No 9 (1 Nov 1964).

Hemant H Dave'- inventor at Applied Research Corp, researcher
at NASA Goddard SFC Laboratory for Extraterrestrial
Physics, personal communication (Sep 1987).

Drake Deming & Michael J Mumma- (83) [op cit Chapter 2].

Arthur D Fisher- (85) [op cit Chapter 2].

C Freed, L C Bradley, R G O'Donnell- "Absolute Frequencies of Lasing Transitions in Seven CO₂ Isotopic Species" DoE contract work by MIT Lincoln Laboratory (no date).

George Gatewood- University of Pittsburgh Allegheny Observatory (personal communication, Apr 1987).

E I Gordon & M G Cohen- "Electro-Optic Diffraction Grating for Light Beam Modulation and Diffraction" IEEE Journal of Quantum Electronics Vol QE-1 No 5 (Aug 1965).

Michael D Griffin- "Spacecraft Systems Engineering" lecture notes at Univ of Maryland (1986).

Gerhard Herzberg- Molecular Spectra and Molecular Structure Vol III (Van Nostrand Reinhold, 1966).

Francis A Jenkins & Harvey E White- Fundamentals of Optics 4th ed (McGraw-Hill, 1976).

Pao-Lo Liu- "Surface Wave Electro-Optic Modulator" High Frequency Optical Communications SPIE Proceedings Vol 716 (1986).

NASA Tech Briefs- "Lightweight Metal Mirrors" p 121 (Fall 1983).

Harvey C Nathanson, Michael C Driver, R Noel Thomas- "Gallium Arsenide Changes the Shape of Airborne Radar" Aerospace America (Apr 1987).

Gerard K O'Neill- The High Frontier (Bantam, 1978).

Optical Industry & Systems Purchasing Directory (Optical Publishing, 1984).

R A Outlaw, W K Peregoy, Gar B Hoflund, Gregory R Corallo-
Electron Stimulated Desorption of Atomic Oxygen from
Silver NASA TP-2668 (Apr 1987).

R A Paquin, H Levenstein, L Altadonna, G Gould- "Advanced
Lightweight Beryllium Optics" Optical Engineering 23(2)
pp 157-8 (Mar/Apr 1984).

Raymond J Roark- Formulas for Stress and Strain 4th ed
(McGraw-Hill, 1965).

Carl Sagan- "Organic Matter in the Solar System: A New Space
Age Paradigm" presented at NASA Space Life Sciences
Symposium: Three Decades of Life Science Research in
Space, Washington DC (21-26 Jun 1987).

Anthony E Siegman- Lasers (University Science Books, 1986).

Bruce A Smith- "Martin Lab Halves Time Needed to Retarget
Space-Based Laser" Aviation Week & Space Technology
(14 Sep 1987).

Jun'ichi Sone, Jaw-Shen Tsai, Hiroyuki Abe- "Picosecond
Josephson Logic Gates for Digital LSIs" in Picosecond
Electronics and Optoelectronics (Mourou, Bloom, Lee
eds) (Springer-Verlag, 1985).

Keto Soosaar- "Precision Space Structures" (Dec 1984) in
[Boyer, 85].

Ronald R Stephens & Richard C Lind- (78) [op cit Chapter 2].

Orazio Svelto (David C Hanna, tr)- Principles of Lasers 2nd ed
(Plenum, 84).

Dietrick E Thomsen- "Big Telescopes on a Roll" Science News
Vol 132 pp 170-1 (12 Sep 1987).

Two-Six Incorporated- "Comparison of Reflection and Figure of
Merit for Enhanced Gold and Enhanced Silver On Si Total
Reflectors at 10.6 Microns" Technical Data Sheet
(Saxonburg PA 16056, no date).

Unreferenced illustration taken in 1973 from high-school
physics text.

A F Whitaker, S A Little, R J Harwell, D B Griner, R F DeHaye,
A T Fromhold Jr- "Orbital Atomic Oxygen Effects on Thermal
Control and Optical Materials: STS-8 Results" NASA Tech
Briefs (April 1987).

Gordon R Woodcock- Space Stations and Platforms (Orbit, 1986).

Amnon Yariv- (75) [op cit Chapter 2].

Paul R Yoder Jr- Opto-Mechanical Systems Design
(Dekker, 1986).

CHAPTER 8

SPACECRAFT CONTROL

Chapter Abstract - Making millions of parts distributed throughout Venus space act together as an optically stiff laser system is the toughest large space structure controls problem yet framed. The fleet achieves it by being one extended, actively intelligent robot. Employing a flexibly parallel organization inspired by vertebrate neurophysiology, the fleet controller uses inertial and relative state sensors to develop motor commands in accordance with its mission plan, executing them via hierarchical actuators (thermal, piezoelectric and electromagnetic) embedded in the fleet hardware. Light delay limits state feedback to 6 Hz. Passive techniques and disturbance avoidance minimize the active control task, and periodic perturbations become predictably familiar to the learning fleet intelligence. Robustness and efficiency increase over time. Advanced star trackers provide the fleet's ultimate pointing reference.

Operating our large optical satellites within the tolerances required by Chapter 7 separates conveniently into two parts: orienting the spacecraft buses properly once they are on station ("coarse" attitude control, addressed in Chapter 9), and then cooperatively micro-adjusting the tilt and focus of all the mirrors through their attachments to those buses (vernier mirror control). The mirror control we need far surpasses any contemporary fine-pointing abilities, due to the extreme precision with which millions of elements, separated by large distances, must work together.

Intelligent Structures

The inevitable flexibility of recently planned kilometer-scale, Large Space Structures (LSS) has necessarily spawned a new field, given the appellation Control Of Space Structures (COSS). Rigid body motion becomes a forgotten dream upon graduating from small satellites to LSS. Over large distances, particularly in jointed structures, even minute and widely distributed strains can result in quite large relative displacements, easily ruining the exacting dimensional stability required by communication facilities. Also, the limited speed of mechanical (acoustic) propagation through materials introduces delays between a load and its distant responses. In general as sizes increase, fundamental vibration frequencies drop, until eventually they intrude into the bands typically used by spacecraft control systems, which can result in unstable oscillation [Herzberg, 84]. Finally, as noted in Chapter 7, environmental forces which would otherwise be small perturbations (like light pressure and gravity gradients) can for LSS constitute major inputs [Szirmay, 79].

Coss is divided into two categories defined by disturbance type [Haftka & Adelman, 85]. Manufacturing and assembly imprecision leave permanent figure deformations which are compounded online by slowly varying ambient disturbance fields, notably temperature. These quasisteady influences are typically neutralized by slowly-applied shape control measures like local thermal actuation. Most effort so far, however, has addressed transient distortions of the structure from applied loads, or Vibration Control (VCOSS) [Aubrun et al, 83]. Even though material damping would eventually attenuate the effects of isolated impulse loads, the weak Coulomb (friction) damping in these sparse structures and the absence of viscous damping in vacuum would in the meantime permit oscillations leading to instability, or simply ruining optical accuracy. Besides, typical spacecraft suffer from continual vibration sources (which we will characterize later) in addition to impulses, so an unattended LSS might never settle down.

Recognizing the need to correct undesired shape changes, much research emphasis is now devoted to controlling actively the internal and overall displacements of LSS using servoelectricity [Atluri, 87], typically by replacing "dumb" structural members with active members whose length or shape can be changed under processed control. The central, simplified closed-loop approach is familiar and completely analogous to the active phase control reviewed in Chapter 2: use sensors to determine the actual system state, combine those data with the desired system state by means of appropriate control laws to develop correction signals, and finally apply those signals to actuators placed within the structure, deforming it willfully to compensate environmental influences. Such mechanisms have, somewhat prematurely, come to be called intelligent structures (a term first applied overenthusiastically to flight vehicles studded with sensors to

predict their fatigue life [Rowe, 86]). As we will soon see, the planetary laser requires true intelligence.

But active control alone is not the best approach. Recent work has begun to develop enhanced methods for passive vibration suppression [Sesak et al, 87], with the goal of optimizing the mix of active and passive techniques designed into a given structure [Simonian, 85]. Three basic types of passive dampers are distinguished by their operation. First, joints with predictable damping properties are arising from new theories of intersurface friction which model stiction [Amos, 85]. Second, discrete dampers arranged along load paths dissipate mechanical energy (as heat) with viscous fluids (dashpots), or internally hysteretic solids (constrained viscoelastic materials), or electromagnetic eddy currents. A third, more sophisticated, approach uses such discrete dampers to absorb vibration energy removed from the plant through spring-mass mechanisms adjustably tuned for certain frequencies. Common in rotorcraft and industrial machinery, these can amplify plant kinematics by up to an order of magnitude, enabling efficient damper "working"; just two tuned dampers, with a total mass less than 10^{-3} that of NASA's Space Station, can increase its structural damping up to $\zeta \approx 0.2$ [Sesak et al, 87].

Although passive devices can greatly reduce the task left to active components, optically accurate LSS still clearly require the quick precision offered by active intervention. Enormous size however introduces as-yet unsolved complications, which have already revealed directions of new research pursuit. First, a continuous medium like a beam has an infinite number of degrees of freedom (DOF), and so cannot be controlled to optical tolerances using any form of discrete classical control theory [Atluri, 87]. This is so because the dimensions of even one member in an assembled structure exceed by many orders of

magnitude the radiation wavelength, so lumping parameters at structural nodes neglects the real disturbances happening between them. Simply put, the structural model resolution in this case is hopelessly coarser than the optical resolution required. Our design avoids much of this problem by requiring in fact the control only of discrete mountings, spaced meters apart, and letting the monolithic mirrors themselves handle finer scales passively. Still, a basic vertex station with almost 230,000 mirrors requires discrete control of at least three times that many actuators (actually many more, as we will see); something more than simple classical control is needed.

The second complication is that using millions of active components assures that stochastic failures will characterize normal operation. LSS work has commenced studying both system degradation given varying degrees of redundancy [Haftka & Adelman, 87], and modelling techniques which can identify rapidly and isolate a failure address to enable compensation [Baruh, 85]. Component reliability represents only one source of temporal plant variation, though. Such long-term effects as material creep, and degradation in the space environment, are sure to produce deviations of at least optical significance (of order nm) from any initially calibrated state. In addition, dynamic properties will change drastically during assembly of any LSS [Szirmay, 79].

The only way to compensate such changes is to make the Active and Passive control (APCOSS) adaptive as well, something not yet possible [Atluri, 87]. A truly intelligent structure will thus monitor and adjust itself to maintain nominal performance over periods ranging from milliseconds to years. Clearly the quicker such adaptiveness can be, the more effective and versatile even normal active operation can become. The stringent requirements of controlling a planetary resonator thus drive both APCOSS and so-called Artificial

Intelligence (AI) to levels far beyond their present states, as we shall see presently.

Processing

Sensors and actuators constitute respectively the receptor nerves and muscles of the robot body. Between them, interpreting incoming information to generate responsive outgoing signals, must operate a brain. Although many schemes are being developed for COSS, standard methods so far are based on Linear Quadratic Gaussian (LQG) theory [Amos, 85]. Defining LQG features will show us just how much more evolution is necessary to meet our needs. LQG theory [Amos, 85]:

"requires a linear model of the structural dynamics, a linear relationship between the forces and input signals of the actuators, a linear transfer matrix relationship between ...sensor outputs...and the actuator input signals, and the minimization of a quadratic 'cost' functional of the dynamic state in the presence of Gaussian white noise as the criterion for the selection of the various coefficients in the transfer functions."

Neither the actuator operation nor noncollocated sensor/actuator relationships can be so categorically linearized for our purpose. Next, COSS for the planetary laser does not simply mean damping vibrations so the structure doesn't tear itself apart; rather it means cancelling all deviations which would exceed the optical tolerances we have specified; the scale of the problem is much finer than that

envisaged for LQG control. Also, although presuming full spectral (white) noise simplifies the theory, it automatically misrepresents a real mechanical system which must exhibit its own unique, non-uniform and changing spectrum. This shortcoming of oversimplification amplifies a final, devastating problem: computational delay. The simple optimization procedure used by LQG theory results in general in nonlinear matrix differential equations which must be integrated numerically to get the control law coefficients. This "is so computationally intensive as to be feasible only for problems with" at most tens of state variables (using vintage 1985 mainframe computers) [Amos, 85].

The common escape, working with reduced order models, fails our needs on several counts. Recall that by mounting rigid mirrors we have already reduced our controlled system DOF from infinitely many to three for each mirror ($3 \times 230,000 = 690,000$ for just the active reflector of just one of the basic vertex stations). This, however, is clearly not what contemporary ACROSS means by a "reduced order model", since it is computationally several orders of magnitude beyond current ability. Order reduction also means that the global stability criteria of classical adaptation theory are inapplicable to contemporary LSS control, because approximate-model errors cannot be reduced arbitrarily [Amos, 85]. So current methods cannot meet our requirement for adaptation over time, either. Finally, as mentioned earlier, a unique feature of LSS transient dynamics is the comparatively long propagation time required for disturbances to pass through the structure. Amos [85] calls this the "propagating wave event" and notes that traditional vibration mod[al] analysis cannot adequately model such disturbances because their "highly localized nature... requires the superposition of large numbers of modes", again a real-time computational nightmare.

Even the embryonic discipline of Controls Structures Interaction (CSI) is thus already too retro-sighted for our needs. By emphasizing the "sub"-systems dichotomy (controls applied post facto to structures, and how those structures interfere with their own control), this view precludes any chance of solving the challenging problem posed by controlling a planetary laser. We must think instead, right from the start, of designing a controlled structure [Amos, 85], one whose active control is embedded in the very fabric of its entire structure, working with and capitalizing on its inherent limitations and subtle responses, rather than trying stupidly to overpower them. Both the traditional structural engineer and the traditional control engineer would demand some evidence that such sophistication is possible. Indeed, the sequel is really a roadmap of inevitable future work, already familiar to the artificial intelligence engineer.

Control Approach

The basic elements of a path to viable real-time processing for the resonator satellites have already been published. With uncanny prescience, Szirmay [79] predicted that a decentralized control system would comprise the ultimate solution to ACOSS. Albeit cautiously, he set forth the essential features of such an approach: simple control of individual structural units, based on local measurements and local actuation, but engaged in "limited communication" with other segments to allow satisfactory overall performance. The advantages he noted are precisely those we need: a "simplified" control system less sensitive to modeling errors, requiring reduced in-flight computation, and permitting easier on-orbit assembly.

Some current work indeed focuses on segmented structures, each passively and actively controlled and all working together hierarchically [Amos, 85]. Atluri [87] has been developing detailed analytical algorithms tailored for onboard, online μ sec control, which sidestep the "big and dumb" numerical difficulties of, for instance, Finite Element Methods (FEM). These (in some cases ad hoc) nonlinear methods can reduce computational complexity by several orders of magnitude, and may lead ultimately to adaptive, truly intelligent structures, capable of perfect VCOSS --- damping out travelling wave vibrations due to local impacts immediately and locally, within one structural unit, before they excite the LSS.

More than any space system yet proposed, the Venus laser "will eventually and inevitably lead to design approaches which go far beyond present practices" [Szirmay, 79]. No control system yet exists capable of coordinating millions of actuators distributed over thousands of kilometers, based on information from millions of sensors equally distributed, under continually changing environmental conditions, with optical precision, despite hardware failures, for years at a time. Such control is, even just quantitatively, so far beyond current missions that it represents qualitatively an untouched realm. Half of any feasibility study consists of determining how its problem could be solved immediately. Necessarily grounded in understood technology, most of this work indeed does just that. But slavishly following that rule in this chapter would dispatch immediately any possibility of planetary lasers.

The other, often more fascinating, half of a feasibility study consists of determining how its problem might be solved in the future. Declaring a feat impossible based on only a primitive understanding of the relevant principles is often embarrassingly reckless, as the history of technology amply

proves. Indeed, suppressing as unobvious the inevitable outcome of current research would be indefensible in this work. Instead, to design the fleet controller we set forth the requirements it must meet, the "existence proof" of an analogous solution, and the real paths of progress likely to result in success, thereby upholding responsible scholarship. The technology outline which follows is thus prudent projection rather than science fiction.

We focus directly on the orbital resonator, since its performance constraints as derived in Chapter 7 are in general the most severe. (The problems of controlling and redirecting the outcoupled beam are not trivial, of course, but they can be derivatively solved once a resonator can be made to produce the beam in the first place.) The orbiting craft are subject to a variety of disturbances which we will catalog and treat fully later in this chapter. Some, like the aerodynamic drag in Venus' tenuous exosphere, are gentle and essentially constant. Others, like the effects of sunlight, are periodic, piecewise continuous and smooth functions of orbital anomaly. Some, like planetary gravity variations, are continuous and smooth over seconds, steeply graded over minutes, and periodic over hours. Yet others, like internal mechanical vibration, are continual but discrete. Finally, those resulting from events like meteoroid impacts are stochastic impulses, and can be quite severe.

The fleet's response to this incessant and multivariate disturbance spectrum is a kind of triage. Those forces too weak or too fast to compromise the mirror segments' infrared performance are ignored. Those which are global but constant are compensated by the bus propulsion and attitude control systems. Those which affect directly and continually the resonator mirrors' ability to sustain lasing are compensated locally by active control. And those too infrequent,

unpredictable and energetic to be compensated at all are ignored also. Most importantly, of those disturbances which do warrant attention by the fleet controller, the periodic ones get characterized in greater detail upon each repetition; the controller learns its response behavior to them so well that it eventually compensates them predictively.

That ability to learn, a cyclically reinforced sequence of analyzing, remembering, predicting, recognizing, and comparing, is essential for three reasons. First, as we shall soon see, predictive evaluation greatly simplifies (makes more feasible) the real-time decision processing required to keep the resonator lasing, by obviating the need to counter every disturbance detail anew. Simply stated, a smart controller is more power- and time-efficient, and more reliable, than a dumb controller. But second, that predictive advantage cannot cheaply be bought with merely rote behavior. Given the variable disturbance field in which our fleet is immersed, a controller incapable of recognizing patterns, adapting to them, and then recognizing the higher-order patterns which disrupt the simpler patterns, would be virtually useless, permitting only sporadic lasing. Truly intelligent learning is strictly necessary for long-term mission success.

Third, the mundane but inescapable fact is that the resonator satellites are so far apart that the nonzero light propagation time between them limits the update rate of system state intelligence. No satellite is more than two ring sectors (60 msec plus processing time) away from any given other, but the state rate is incontrovertibly limited to 6.7 Hz by the 150 msec (plus relay processing time) light delay all the way around the ring. That is, the effect on the laser of any satellite's actions cannot be known by it sooner than 0.15 s later; feedback loop closure is thus severely rate-limited. This is not to say that the satellites cannot take any action

at faster rates; indeed Chapter 7 showed that they all must. While dither control of a laboratory resonator at rates slower than 6 Hz can be practical, remember that our resonator mirrors orbit independently in different local disturbance climates, at over 6.5 km/s. The fleet controller must have an excellent idea of the immediate consequences of its actuation, since it can only verify the results 150 msec later. Learned predictive control is therefore vital.

Nonperiodic disturbances in the range requiring active compensation obviously cannot in general be predicted, and so present the controller with its toughest challenge. As much as possible, we avoid such disturbances by design. That is, to assuage the controller's job and thus enhance its success, we motivate many subsystem selection choices according to how well they suppress, eliminate or preclude random disturbance forcing of the spacecraft. The theme even of sacrificing lightness and efficiency for the sake of dynamically quiet operation recurs often in Chapter 9. Sidestepping that dual goal of conventional spacecraft design -- lightness and efficiency -- may seem rather cavalier. But since our goal here is to navigate a path to feasibility for a system with remarkable needs, we must consider the sacrifice a price well paid. The cost analysis of Chapter 11 will justify the tradeoff in any case.

The controller is a distributed, hierarchical brain which generates sensory data at its lowest level, processes them to extract increasingly important, general features in progressively higher levels, evaluates that encoded state performance with respect to its high-level mission standard, and generates governing system directives that, once translated into specific commands by succeeding lower levels, drive appropriate motor systems. As response patterns repeat more often, their control is shifted to lower, more autonomic

levels, shortening response time and freeing the higher levels for more efficient general analysis. That opportunity to address longer-term patterns then allows better prediction; as even the longest periods are well-characterized, the again-liberated excess high-level capacity transmogrifies finally into a highly redundant (and therefore reliable), streamlined top-down command network. Performance verification several times each second (for years) continues to hone the brain over its mission life. Ultimately able to manipulate its distributed body with coordinated assurance, the brain can then easily control the fleet into sustained lasing for long times in the comparatively benign, even boringly familiar, disturbance environment around Venus.

During normal operation, the highest processing level (distributed among all the resonator satellites) develops a continuously revised master plan of how the resonant laser beam should be. Each satellite aspires physically to meet its ideal place within that standard scheme. Many voting, tuned interferometric sensors distributed across each craft enable it to "ride" the resonant cavity wave by establishing a reference plane intersecting the spatially coherent beam with a phase determined by the master plan. The mirror segments, organized into neighborhood groups linked in turn under the control of increasingly inclusive domains of the entire reflective surface, match the desired reference plane. Highly cross-checked inter-domain sensors enable the individual segments to act as one, becoming in fact the reference plane and shifting as the master plan shifts it, according to environmental disturbances. Thus by using a hierarchical sensory and command structure, the fleet can compensate quickly enough for environmental changes to maintain a continuous-wave laser beam.

Clearly the ability to recognize and remember important patterns, generalize in real-time, reallocate its own circuitry

according to practiced skills, and ride herd simultaneously on millions of actuators, comprises a tall order by current standards of automation. To readers familiar with the catchwords of modern computing, the obvious processing approach is a "massively parallel" one. But we need to examine why, and to see that much more than just parallelism is necessary. To pursue these topics, and to demonstrate that the kind of control our fleet needs is not without precedent, we will look briefly into an operational analog: vertebrate neurophysiology. Studying a system which already for millions of years has exhibited sensory intelligence, adaptive learning, flexible redundancy and exquisite simultaneous control similar to that required to run the planetary laser, and is vastly different from conventional computers, provides valuable insight into what we expect the working fleet controller must be. Except where noted specifically in the sequel, the reference for this review is Kent [81].

The basic processing units of the brain are its neurons. The human brain starts with about 10^{13} neurons, of which up to 85 % are killed in normal infant development [Churchland, 86]. This grim selection establishes a central feature of the organic brain's operation: like an artist who draws with both pencil and eraser, it works as much by inhibiting potential as it does by building complexity. Each neuron is a marvelously versatile "gate" able to perform any of the logic operations familiar to circuit designers. It receives input stimulation and inhibition from other neurons on its cell body and along its dendrites, and delivers pulse-coded-intensity output stimulation to others with its axons, ultimately controlling muscles and glands. Axons and dendrites extend throughout the animal's brain and body, gathered in dense bundles, to target specific sites. The stimulation itself is electrical, mediated rather slowly by chemical transport across the synapses between neurons. The human brain draws about 25 W.

Soberingly unlike contemporary manufactured gates, an average neuron receives about 10^4 inputs. Each such signal line is spatially encoded by its synapse placement, and temporally encodes intensity by its firing rate. The brain's operation is nonsynchronous; results trickle among the processing echelons without being clocked. A neuron fires like a Schmitt trigger when the analog sum of its inputs exceeds some threshold, and then resets. But because feedback stimulation, input weighting and threshold biasing are all variable for neurons, their logical function is most aptly described as an ALMOST gate. At the cost of some imprecision, this approach buys speed and flexibility; entire sensory networks, for example, may be tuned by other, control, regions through selective inhibition. The fail-safe operating condition is one of general excitation (driven by the lower reticular formation), out of which selective inhibition (from the higher cortex) "carves" an activity pattern attending to the analysis being performed. Consciousness seems to be the feedback maintenance of this motivation, and a thought is "the temporal sequence of spatially ordered events in the brain."

The problem of how the brain commits neural activity patterns to long-term memory is still rather intractable, but the massive interconnection implied above argues as much for extreme redundancy as it does for widespread feedback control. Indeed, neurons are unique among cells in not being replaced when they die, which they do at the rate of about 10^3 /hr. The brain can "run a relatively constant program in a varying supply of parts", and clearly encodes perceptions as distributed patterns in cell populations rather than in dedicated and vulnerable single active lines. Transient activity leaves a stabilizing trace in the changing brain; with repetition or associative reinforcement the trace eventually becomes structure. This interaction between the brain's

patterns and its environment "encompasses time scales ranging from tenths of a second to hundreds of millions of years" [Changeux, 85].

The visual system serves as a good paradigm for our fleet sensors because of its simultaneous dependence both on detail and collective meaning. Several characteristics enable the brain to process sensory data efficiently. It updates its analysis of the complete visual field (about 10^7 light receptors in each retina) at about 10 Hz. Since neurons cannot fire faster than at about 1 kHz, that visual analysis is done within a mere 100 information-processing steps [Churchland, 86]. Parallel analysis is obviously extremely useful. The feature-extractor mechanism which reduces so much data into a manipulable form apparently performs real-time frequency analysis of the visual field, so it works immediately on patterns rather than raw form. In general, evolution has favored neurosystems which match complex sensors to "simple" brains; sensor preprocessing is obviously also extremely useful.

Information moving up to the higher levels is encoded in a variable-length and variably-accessible word, so that each level's analysis is added on in parallel, and any functional unit can access any bit in the code at any time, even simultaneously with other units. By thus avoiding analysis bottlenecks, such nonsynchronous operation too is extremely useful. Intermediate results can motivate sensory feature-extraction tuning and motor responses even as they are analyzed by higher levels. Because the original data are transmitted unaltered along with the analyses they accumulate in the variable word, even the highest centers can use them, and then reach down practically to the lowest level to effect concerted control over specifics.

Most often, however, the cortex transfers both symbolic and actual (motor) sequences it has learned to subcortical control, avoiding "the need to reason out problems anew on each occurrence... Reasoning is...fundamentally the same process as the design of complex goal-directed motor behavior, and in fact supported by the same hardware." Reasoning and memory are enhanced for heuristic thinking by the "fuzzy" address permitted with neuronal ALMOST gates. Associative thinking, guaranteed by the parallel interconnected neuronal structure, can solve problems lacking a single correct solution, not perfectly, but adequately, quickly, and under changing circumstances of input and degradation. The price, of course, is a memory which is inexact, and an evolutionarily precluded ability to perform extended and precise symbolic analyses. (For those jobs our organic brains have developed the serial binary devices we normally think of as computers.)

Now we can with some real basis imagine a model controller for the planetary laser which combines useful features of both familiar computer technology and less-familiar brain "technology". Linked by modulated light across the extent of each satellite and the vast separation between them, it must be a densely interconnected and massively parallel web of optical gates, each somehow flexible enough to be rebased on-line and "develop" new connections with other distant gates, and versatile enough to perform both the "hard" and "fuzzy" logic functions required by precise motor control and adaptive thinking respectively. It has to generate with tuned feature extractors useful and compact representations of what its sensors tell it, accumulating analyses of those data and storing them accurately but accessibly for later use. Its conscious thoughts evolve a complex command regimen for the fleet actuators, testing it and improving it over time, and finally delegating its routines to autonomic processors. Higher levels then devote their resources to building an

indissolubly redundant control network embodying what the brain has learned about its operational environment at Venus.

Lest the skeptical reader object to positing the manufacture (as a spacecraft subsystem!) of a brain of vertebrate complexity, we point out some salient distinctions and current events. First, the fleet controller shares with even primitive industrial robots the trait that it need only solve one kind of problem. Its intelligence need not wrestle with altogether fresh situations as a matter of course; really the extent of this brain's novelty is limited to its adaptive, improving control of a rather constrained situation, and its physical distribution around a planet. A human brain with much experience driving a car, for instance, rarely devotes higher cortical levels to the task, but rather proceeds "in a stimulus-bound, feedback-controlled mode of operation, which does not differ in principle from that of a lizard approaching food" [Kent, 81]. The image of a lizard brain is perhaps an appropriate one for keeping the fleet controller in perspective. While quite complex, and able to control a mechanism no human ever could, it in fact is so much simpler and limited than a human brain that it does not even require limbic (emotional) motivational systems. In a very real way, the itinerant fleet maintenance robots of Chapter 9, not the laser itself, require the most challenging heuristic intelligence that the fleet controller must demonstrate.

Some initial steps have already occurred along paths leading to the type of artificial intelligence required by our fleet. Logic switching techniques for optical computing are being developed [Neff, 86] precisely because they hold the potential for high speed (much higher than biological nerves, incidentally), parallelism, and dense interconnection, and because they interface so simply with optical sensors [Pisacane, 87]. Meanwhile parallel electronic computer

architectures are being developed, producing great strides in speed and ability for hardware matched to the problems they solve. An example is NASA's Massively Parallel Processor (MPP) which, albeit only a two dimensional processing array, has already enabled investigations into entirely new ways of solving analytical problems representable for centuries only by differential equations [Wolfram, 86].

Other studies using the MPP have direct relevance for developing competent artificial intelligence. Simulated "neural nets" [NASA TB, 88] run on such a device can model certain brain functions, including gradualism and modification-based learning. Even with simple threshold "neurons" and threshold "synapses" to connect them, a few rules for repetition learning can enable impressive soft programming ("the indirect control of the evolution of the system by the environment") [Hastings & Waner, 86]. The MPP implementation of such studies runs much faster with shorter programs than serial versions and, significantly, does not slow down as the number of firing neurons increases. If such "architectures are especially appropriate and useful for neural network" simulations, then we can expect their derivative technologies to transcend just simulation and become truly useful in their own right.

Finally, we note that as part of an enormous continuing effort to learn more about biological intelligence, fabricated circuits are being connected directly to neurons, to monitor and eventually control their activity [Miller, 86]. Thus the boundaries between natural and artificial intelligence blur relentlessly, dissolved from all sides as many researchers pursue goals embodied by our fleet controller. Given the state, rate and directions of current work in advanced computing technologies, the inarguable existence of lizards which approach food generation after generation, and the

comparatively modest (even if extensive) complexity required by our controlling several million actuators cooperatively in a repeating environment, we must conclude that building and operating a brain to run the planetary laser will indeed be feasible in the future.

Perturbations

Next we catalog the perturbations with which the fleet controller must cope. While this section will make Venus space seem Bosch-like by the extent and detail of its hellishness, we should start by pointing out that in fact, the vacuum and microgravity of space provide in general the most optically disturbance-free environment attainable. True, we must use great effort to capitalize on its attributes; but only the relatively benign, "weatherless", contactless surroundings of space make a planetary-scale laser thinkable at all. Although they overlap ambiguously when affecting large satellites, we separate the relevant mechanical disturbance sources into global (field) and local categories for clarity, and begin with the former.

Gravity variations constitute serious disturbances. Four classes concern us [Wertz, 84]: Venus' own non-spherical potential, its gravity gradient, tugs due to other masses, and relativistic effects. We visited briefly in Chapter 6 the gross station-keeping effects on our resonator of Venus' non-spherical field. Appendix A8-1 confirms our earlier result expecting maximum non-Keplerian radial excursions smaller than 2 km over minutes, resulting from accelerations of about 0.05 m/s^2 , and maximum tangential excursions smaller than

700 m resulting from accelerations of about 0.015 m/s^2 . The fleet response to these undesirable motions is high-frequency mirror actuation (both mode-hopping focus as explained in Chapter 7, and pointing tilt), monitored by the cavity phase sensors and mediated predictively by the fleet controller after thousands of orbits overflying the same terrain.

The gravity gradient experienced by an extended satellite whose extremities occupy different radial positions relative to the planet is insistent enough to stabilize the attitude of even small satellites like the LDEF currently orbiting Earth. Since the gravity gradient attenuates quickly with increasing altitude it is irrelevant for L1 and L2 Stations. Furthermore, being slow-acting (about 10^{-5} Hz [Aubrun et al, 83]), it can be ignored by mirror control. As indicated in Chapter 7 though, the gravity gradient does introduce severe, constant attitude torques for large satellites which do not fly principal-axis oriented, a description fitting directly the large satellites of Station 1. Appendix A8-2 assesses this gravity gradient penalty to show why "expendable" solutions are not practical for huge spacecraft.

Instead our fleet relies on what we can call a "structural", or passive, solution to the constant, large gravity gradient torque. For instance, all of 1α 's torque and one component of 1β 's can cancel each other to first order if the two reflector craft are braced by interlinking structure. The other component of 1β 's torque, and the torques of 1γ , can be produced passively and constantly by tether tension through a CM-offset attachment. Thus at the cost of extra (mostly passive) structure and mass, configured carefully, we can avoid the tremendous logistics penalty of literally thousands of engines and thousands of metric tonnes of propellant expended every year. Of course, inert counter-mass could be distributed about the spacecraft buses so as to make them inertially

principal-axis oriented despite their geometrically asymmetrical attitude. But we shall see momentarily another, surprising need for intercraft structure at Station 1.

Distant massive objects in the solar system can affect orbital mechanics noticeably over long times. Current efforts to detect Jupiter-size extrasolar planets (around other stars), for instance, take advantage of the slight stellar wobble such an orbiting body induces. And Venus itself is gravitationally locked with Earth, keeping always the same side facing us at its closest approach [Hunten et al, 83]. The largest and most variable solar system gravity effect at Venus results from the sun, however. Appendix A8-3 calculates a maximum resulting acceleration of about $1.6 \mu\text{m/s}^2$ on the resonator satellites. Tangential effects (which would change the orbit energy and hence its size) cancel continually, but the radial effect (which pulls the orbit's shape out of circular) accumulates continually and must be propulsively counteracted for the laser to work.

The surprising corollary for Station 1 is that its component craft are so heavy, and so close together, that they attract each other rather strongly. Appendix A8-3 goes on to estimate the specific perturbing force, from one satellite at the location of another, as roughly 40 times larger than the solar effect! Meeting this propulsively for the three large reflector satellites (whose total experienced gravitational force naturally is proportional to their own enormous mass) would be wasteful. Again, a relatively small amount of compressive structure to brace the craft apart could do the job at a fraction of the systems penalty. Albeit geometrically more cumbersome, the structural solution is unquestionably more elegant (we should expect that, with more iteration of the reference design used in this study, Station 1 would probably evolve into a single, more compact satellite).

Relativistic effects are usually ignored as insignificant sources of orbital error [Wertz, 84], but they are measurable in certain cases; Mercury's perihelion changes by about 524 nrad/orbit, explainable by general relativity (GR). For our purposes we note that the resonator orbital velocity of 6520 m/s is $2.17(10^{-5})$ the speed of light, over ten orders of magnitude larger a fraction than the ratio of our allowable uncompensated mirror excursions (62 nm) to the cavity length of $45(10^6)$ m. We take this to indicate that GR effects must be accounted for in the actual operation of the planetary laser, although since such compensation lies easily within the actuation bounds already required by other constraints, we will not pursue it here.

A separate, gravitationally-amplified, perturbation designed into the fleet formation is non-Keplerian motion. Any departure from the gravitationally-defined orbit a satellite would normally follow requires energy expenditure to achieve. Station 1 is not a simple resonator vertex, but rather a distributed constellation serving several functions outlined in Chapter 7. The craft must, however, travel together in formation as though they were located jointly on the same point. We have already arranged the Station 1 geometry to minimize secular forces, first by positioning 1α and 1β so that their individual CMs ride on the orbit circumscribing the resonator pentagon. Then we connected 1δ and 1ϵ with a short tether so that their combined CM also rides on that same path. But 1γ is constrained by optical (diffraction) parameters and its size to ride -- in formation -- both at a slightly lower altitude than, and slightly out-of-plane compared to, the nominal resonator orbit. Even though the discrepancies amount to only hundreds of meters, their constant nature adds up to a sizable penalty for a heavy satellite, as Appendix A8-4 shows. The best solution, familiar by now, is the passive one of

connecting 1Y structurally, using tethers and outriggers, to counterbalances and to the other large craft of Station 1. By providing intercraft load paths, we do indeed locate them jointly at a single point, as far as the orbital mechanics is concerned.

Another global disturbance affecting laser performance is the ambient magnetic field through which the resonator satellites fly. Such fields affect satellite attitude by torquing onboard dipoles. Such dipoles are of two types. Residual magnetic moments result from the permanent signatures of magnetized components. We avoid these altogether by building the fleet in general out of paramagnetic and diamagnetic materials. Such ferromagnetic materials as we do specify (like the permanent magnets in EM actuators and the AMCD rims) are contained in housings providing flux return paths to minimize their distant free-space magnetic moment. Induced magnetic moments are typically the more severe problem for high-power spacecraft, resulting from onboard current loops. They are avoided in detail by proper power and electrical signal system choices (coaxial or twisted conductors) and layout (balancing the net area-current product of loops with opposite sense).

Two magnetic field contributions concern us at Venus. The intrinsic Venusian planetary field is often taken to be zero, as noted in Chapter 6, but only by comparison with the Earth's field. Adapting surface data [Hartmann, 83] to our altitude, we ascribe a 19 nT strength to the planetary field, which dominates at least antisolar portions of the resonator orbit. This field is so weak, however, that the 5 nT interplanetary magnetic field, generated by the sun and carried by the solar wind [Wertz, 84], produces a subsolar bow shock virtually coinciding with our orbital altitude [Smirnov et al, 80]. Thus the resonator satellites pass through the turbulent

magnetosheath, with its fluctuating field strengths and directions, during a substantial fraction of each orbit. We expect rapid magnetic shifts to have essentially no effect on the enormous, magnetically neutral masses of our fleet satellites.

We might think it possible though, given better characterization of the interplanetary field's sectorized and transient structure [Wertz, 84], to torque against it using dedicated current loops for angular momentum control or desaturation (Chapter 9). Appendix 8-2 noted in passing, though, the impracticality of effective magnetic torquing for our satellites, even using the stronger Venusian field. A related orbital perturbation is drag due to onboard eddy currents induced in the spacecraft's conducting materials by its passage through ambient magnetic fields. Since even the Venusian field is three orders of magnitude smaller than the geomagnetic field for orbits where this effect is considered insignificant [Wertz, 84], we will ignore its effect here..

The solar wind is a tenuous plasma formed from coronal gas ejected by the sun at high velocities (about 300 km/s) [Wertz, 84]. It is barely characterized at distances other than 1 AU from the sun, and not at all outside the ecliptic plane. High-velocity streams occur sporadically, doubling the "quiet" velocity for a few days at a time. With a mean radial integrated momentum flux of only 4.4 nN/m^2 (adjusted to Venus' orbit from Wertz's Earth data) and a mean non-radial value three orders of magnitude smaller than that, the solar wind itself is a minor perturbation source compared to, for example, solar radiation pressure. Any hypothetical, fast, small-scale variations in the wind, even if they could disturb local mirror positions, could certainly be cancelled by the 6 Hz closed-loop mirror control.

Another "wind" force resulting from the integrated flux of individual particles striking the orbiting satellites is drag from Venus' exosphere. Typically, atmospheric drag becomes less important than solar radiation pressure above altitudes of order 500 km at Earth, although long-term perturbations do result at altitudes up to 1000 km [Wertz, 84]. No detailed information yet exists concerning drag effects of Venus' upper exosphere, although diurnal swelling due to the planet's slow rotation causes a factor of ten departure from the simple exponential density model based on hydrostatic equilibrium, as noted in Chapter 6. Appendix A8-5 offers an excessive upper bound on atmospheric drag at our 1589 km orbital altitude, resulting in a worst-case net force easily made up propulsively. For our reference configurations, we will consider no offset between the CM and center of pressure (CP), thus avoiding aerodynamic torques on the satellites.

Our final global perturbation is radiation pressure, the result of photon momentum exchange. Here again, we will consider no CP/CM offset, precluding radiation-induced torques (these will in fact occur at the terminator crossings, but we expect their effect to be much less severe than the simultaneous direct transient impulses considered shortly). The four radiative contributions to propulsive station-keeping are from the laser beam itself, direct sunlight, sunlight reflected from Venus, and IR emission from Venus. The laser's intracavity power density, at 11 W/m^2 , is trivial compared to the 2660 W/m^2 solar constant at Venus. The intermediate beam (the output beam focused down to a 10 m diameter) however, which impinges at full strength on the 1δ and 1ε craft, has a power density of 2300 W/m^2 (Appendix A8-6), practically as strong as local sunlight for those small mirrors. Appendix A8-6 also shows that the dayside reflectance of Venus contributes 1225 W/m^2 , of the same order as the solar value; the planet's 103 W/m^2 thermal emission is an order of

magnitude weaker, albeit constant throughout the entire orbit, day and night. Appendix A8-6 goes on to show the combined effect of these radiation sources on stationkeeping throughout the fleet. Compensation must be propulsive.

The first of our local disturbances derives directly from this solar radiation also. The legendary abruptness of orbital sunrises and sunsets generates a substantial transient disturbance which sweeps across each of our large satellites in turn as they cross the orbital terminators. Appendix A8-7 begins by evaluating the percentage of each orbit that the satellites suffer from solar pressure, then shows the maximum force density experienced in the fleet (by the 1 α satellite) to be 16 $\mu\text{N}/\text{m}^2$, switched on and off each orbit within just 0.25 s. The terminator sweeping across the satellite brings with it a differential impulse load which must be compensated as it occurs by the active structure if lasing is to continue uninterrupted.

Sunlight, reflected light, and planetary thermal emission not only push on the spacecraft, they heat them up. We consider thermal control techniques in Chapter 9, but here we note that thermal distortions occurring within the structure constitute local geometrical perturbations which must be accurately measured and compensated. The job is made easier both by its repetition and by thermal inertia. A by-product of active cavity control is that all the craft enter and leave sunlight in exactly the same orientation every time, for thousands of essentially identical orbits. The controller thus learns in detail the dynamic fleet response to the transient load, simplifying its real-time control effort with each corroborative orbit. And because thermal effects occur comparatively slowly even when their source is applied instantaneously, sophisticated compensation is quite feasible.

Our fleet active structure detailed in Chapter 9 performs the actual adjustments.

The last external source of local disturbances we will consider derives from the continual rain of meteoroids our fleet satellites feel. We already expect (from Appendix A7-2) the particle flux for both the resonator satellites and the libration station satellites to be essentially the same. Natural debris is assumed to have an asteroidal or cometary origin [Howell, 86]. Sporadic meteoroids are those occurring with a random orbit distribution; their speed distribution peaks at about 15 km/s. Meteoroid streams or showers follow closely matched orbits, exhibiting speeds up to 70 km/s at Earth's orbit. Such showers can increase the average cumulative sporadic flux by factors approaching 10, for several days.

The flux of these "very fast moving and hence invisible as well as unavoidable" objects depends on what size we choose to pay attention to [Woodcock, 86]. Based on a composite statistical model, we may presume roughly one impact per square kilometer of spacecraft per year by particles larger than 1 gm. Alternatively, we may expect of order 10^5 impacts by particles larger than 10^{-4} gm on the same area during the same time. That is, about a dozen objects, each delivering perhaps 20 J as an impulse load to some point location, will hit each square kilometer of the fleet hardware every hour during "quiet" sporadic times. During intense showers, that might become a hundred objects each delivering about 250 J, every hour, to each square kilometer. Without a detailed design and simulation in hand, all we can say is that the vibrations resulting from small impacts get damped out quickly and locally by the active structure. The potentially disruptive larger but rarer impacts might result in momentary laser interruption of a control neighborhood until its large

amplitude vibration could be brought under control. In some cases replacing damaged components will be required.

Internally generated disturbance forces will also disrupt precise mirror positioning. These are of two main types: housekeeping and operational. Standard onboard housekeeping disturbances like attitude-control bearing noise and imbalances, coolant turbulence, and other pump and motor vibrations, typically occupy the frequency range between 10 Hz and 1 kHz [Aubrun et al, 83]. Our approach to dealing with these familiar spacecraft nuisances is largely to avoid them through careful subsystem selection. Chapter 9 shows in detail just how much this goal of minimizing optical disturbances can drive the rest of the spacecraft design. We choose an actively vibrationless attitude control method; we avoid moving interfaces of practically all kinds; we employ a continuous, smooth propulsion system lacking moving parts or turbulence; we choose a power plant with no moving parts, cooled passively by capillary flow. Eliminating conventional sources of spacecraft vibration makes it really feasible for the active structure to cancel those remaining vibrations not arising from housekeeping functions.

Minor disturbance forces result from the itinerant maintenance robots which continually attend to the fleet's minor repairs and component replacement. Naturally they are guided by the fleet controller to work carefully, introducing to the structure only the smallest-amplitude forcing which can accomplish the job at hand. Such on-line maintenance work should interfere with the optical performance of at most a small neighborhood, leaving the vast majority of mirrors working normally. People are not permitted to approach the craft.

The last perturbation source, an operational one, is at once the most obvious and the most unavoidable for this mission: the control actuators themselves. Although kHz actuation makes possible the continuous operation of a planetary laser, it simultaneously represents a vibration source which must in turn be compensated. Every motion performed on one of the resonator satellites propagates in some way throughout its entirety, and therefore affects optically the entire ring dispersed around the planet. In an extended structure layered with millions of actuators, we can expect the heaviest control burden to result from the need for those actuators to cancel the structural noise caused by each other. Considering the complexity of this linked problem enables us most effectively to grasp the qualitative challenge of active fleet control for the planetary laser.

Actuators

With this section we begin to evaluate and select specific subsystems capable of executing the control we now envision. Our ultimate actuation goal is to make all the mirrors on a satellite match its reference plane, all the time. The individual mirror segments thus constitute the "payload"; each is mounted to the spacecraft bus at three points. An ability to move those points independently toward or away from the bus allows the three degrees of freedom (focus, and tilt about two axes) which are both necessary and sufficient for micro-aiming.

But it is important to remember that the bus provides the reaction "ground" for all such mirror actuation. Any relative motion between bus and mirror does not move the mirror

absolutely; rather, it displaces both mirror and bus with respect to their mutual center of mass, by amounts inversely proportional to their own masses. The mirrors have been designed (Chapter 7) to move as rigid bodies (to optical accuracy), but the bus, because of its huge distributed inertia force, will tend to react the actuator work with internal, and particularly local, strain energy. Preventing the resulting displacements from interfering with other mirrors' tuning, as well as compensating structural motion resulting from external disturbances, requires additional actuators distributed throughout the bus structure.

Actuators are most easily classified by their operating principles. Before surveying the menu, we should clarify some common terminology useful throughout this section and the next. Resolution refers to the smallest discrete increment with which an actuator can move, or the smallest division of the measurand (property being measured) which a sensor can discern. Accuracy, however, depends on the sum of all system errors and is thus a measure of repeatable performance [Burleigh, 87]. All the motions required in our fleet can be reduced to translation, so we will concentrate on translating actuators.

First are mechanical linkages, based on gears and levers or screws. These are capable of high monotonic speeds over long travel ranges, and can be designed for arbitrary force levels. Various (and expensive) precision refinements can achieve resolutions all the way down to a few tens of nm [Lansing, 84], albeit with strokes only a fraction of a cm. However, because of friction, linkages exhibit two severe liabilities for remote optical use: wear and backlash. The space-tribological problem presented by millions of delicate components operating reliably, as they wear out mechanically, is clearly great. Backlash, the dead band upon travel reversal caused by the dimensional clearances necessary in mechanisms,

limits quick accuracy. No systems in our fleet use such linkages.

Next are fluidic actuators, such as hydraulic or compressed gas pistons, which can also achieve large forces. Such elements have been used successfully in demonstrated structural actuators for LSS [Hoehne, 84] and as low-force springs in precision metrical structures [Yager, 78], but a system requiring moving seals seems to invite more trouble for longterm reliability even than one relying on linkages, so we avoid them also.

We have already discussed passive proof-mass VCOSS dampers; the same principle can be used to make tuned active dampers, driven for instance by the EM actuators covered later. We dispatch these compound devices quickly by noting that, not only do they depend on linkages and/or fluidics to work, they are primarily useful for attenuating motions so large as to be unallowable in our fleet in the first place. The largest local motions our craft experience result from the severe meteoroid impact events we have decided to ignore operationally.

Thermal actuators work by controlling accurately the temperature of a well-characterized solid [Haftka & Adelman, 85] or fluid [NASA TB, 8709] inserted in or applied to the structure. This demonstrated approach is extremely elegant for a number of reasons. First, the subtlety of the effect is easily controlled through material selection and operating temperature range. The coefficient of thermal expansion of aluminum, for example, is two orders of magnitude greater than that of graphite/epoxies. Second, resolution is a function of noise in the thermal source (if electrical, then of the power conditioning), and thermal sensors can be made so fine (as we will see later) as not to compromise system accuracy. Third, thermal strain can just as easily be used to control bending

(via the bimetallic effect) as to control extension. Finally, thermal actuators operate smoothly; they are dynamically quiet. Our active structural members all incorporate closed-loop series thermal actuators to set their length biases (Chapter 9).

Although useful for figure control, thermal actuators are too slow for VCOSS. Memory metal devices, however, have been operated as fast as 5 Hz. Typically made of nitinol (a nickel titanium alloy), the temperature of whose martensitic phase transition can be fixed anywhere between 70 K and 370 K, these reliable actuators feature extremely high specific force. Materially stable, they can be electro-resistively operated, and configured to produce substantial displacements. For example, sub-mm diameter nitinol wires consuming a few watts can exert forces of tens of N and deflections of tens of mm [Studer et al, 86]. Applications would seem generally restricted to binary positioning, though, as the martensitic transition itself is a step function; our fleet uses such devices only for minor jobs like operating sensor telescope covers.

The literature regards piezoelectric (PZ) devices as probably the most versatile and promising for COSS, and their development has been extensive. Some anisotropic materials are so mechanically sensitive that distorting them even slightly, which changes their atomic spacing, produces a large potential difference. Conversely, applying a large potential difference in the "poling field" direction produces expansion or contraction normal to it, depending on polarity. This ability of PZ materials to work both as sensors and actuators, with resolution limited only by the power supply noise [Burleigh, 87], has spurred hopes of truly simple active structures [Atluri, 87] with superior redundancy and reliability, and almost perfect (adjacent) collocation [Fanson & Chen, 87].

Furthermore, PZ materials can operate at high oscillation frequencies, and be formed in any shape: extenders function to produce series axial loads, and if stacked amplify either force or stroke; monomorphs bend structural elements to which they are bonded (much like bimetallic thermal actuators), with the distortion doubled if adjacent oppositely poled layers are arranged as bimorphs [Studer et al, 86]. PZ actuators of various types and sizes are commonly used for mirror control in laboratory and commercial tuned lasers.

The two general PZ material classes are ceramic, represented by Lead Zirconium Titanate (PZT), and polymeric, such as Polyvinylidene Fluoride (PVDF). PZT is a hundred times stiffer, over four times denser, and can generate over four times the specific mechanical motion (m/V) of PVDF. However, PVDF can withstand electric fields 30 times greater, so that larger absolute motions can be achieved; with electrical resistivity seven orders of magnitude greater than PZT, it also operates much more efficiently. Two identically sized (2" x 0.375" x 0.02 ") test samples demonstrate, with a 100 V source and no load: the PZT, with a natural frequency of 140 Hz, deflected 184 μm , producing 59 mN and drawing 66 mW; the PVDF had a natural frequency of 42 Hz and deflected 45 μm , producing 0.31 mN but drawing only 1 nW [Studer et al, 86].

PZ actuators presently commercially available invariably use PZT; knowing that polymer materials outgas in vacuum, and degrade most in the space environment, we should prefer ceramics for our use (but realize that the potential for polymeric PZ materials is great). State-of-the-art units use 150 V supplies, have strokes up to 100 μm , and handle loads up to over 700 N. A typical commercially available power supply with a noise level of mV can produce resolutions on the order of a few nm, and dithering (high frequency

actuation) as fast as kHz. A patented linear motor uses sequential PZ clamping to move arbitrary distances with sub- μm resolution at up to mm/sec [Burleigh, 87]. Reportedly SDIO has tested these in radiation environments for over two years with no degradation. The limitations of PZT are that it is a "fragile ceramic" susceptible to damage from mishandling, impact, and arcing (not usually a problem in hard vacuum), whose operation is ruined by temperatures above the material's Curie temperature (although some PZ materials can be poled as high as 470 - 570 K). Nonlinearity is as high as 5 %, and hysteresis three to four times that; closed-loop operation with accurate sensors is therefore essential.

Our fleet makes extensive use of PZ actuators. Applied as multimorph films to active structural members, they control member shape and dynamic performance. Variations of PZ linear motors allow low-power, precision positioning over large strokes with positive mechanical locking between maneuvers, for systems like the AMCD suspensor/drive station mounts, telescopic telemetry and star tracker mounts, and the Rings' secondary mirror pivots.

Our final actuator category is electromagnetic. As might be expected, this alternative technology, akin to loudspeaker transducers, is extremely versatile. Large actuators, with strokes of tens of cm and exerting hundreds of N of force, can be made although they are rather power-hungry (an efficient linear actuator with 1 mm stroke and 45 N force consumes 30 W), and heavy (1 kg) [Studer et al, 86]. On the other hand, more subtle ElectroMechanical Translators (EMTs) have gained competitive favor for precise micro-positioning. One small EMT designed specifically for CO₂ laser dither stabilization consumes at most 1 W to effect smooth optical component positioning with nm resolution over a 60 μm stroke at speeds between zero and hundreds of $\mu\text{m}/\text{sec}$, at hundreds of Hz.

[Sielmann & Balsarowicz, 84]. It can be used "to bring an object to a certain position and hold it there or to make the object oscillate around a certain position with various frequencies and amplitudes."

Electromagnetic actuator technology continues to improve, partly from design (like a linear-force actuator with dual parabolic windings, which achieves 1 N with 2.7 W [Lange & Holzach, 85]) and partly from new hard (permanent) magnet development. Alnico magnets still boast the highest flux density, but the new cheaper Neodymium-Iron-Boron (Neo-Iron) magnets exhibit the highest coercivity [Studer et al, 86] (coercivity is the demagnetization requirement, so these are more permanent and can be used without weakening in powerful electromagnetic fields). Advanced Samarium-Cobalt (SmCo_5) magnets with low oxygen contamination promise yet higher strengths [NASA TB, 8707/8a], making them obvious candidates for orbital manufacture. Another magnet material achieving 97 % of its theoretically high coercivity when manufactured in microgravity is a Bi/MnBi alloy [NASA TB, 83]. Finally, while the necessary wire length and insulation of EM coils introduces reliability and weight penalties over PZ systems, their resistive (rather than capacitive) nature means that continued unpredictable advances in superconducting technology might make them clear favorites based on power consumption.

EM devices serve as the final, payload actuation stage in the motor hierarchies throughout our fleet. In many ways, carefully designed EM actuators represent the optimal soft mount [Laskin & Sirlin, 86] for space use; we will refer to such actuators as "space bearings". By eliminating direct mechanical contact, the space bearing attenuates high-frequency vibrations automatically; if in addition its electromagnetic grip is actively controlled to ignore specific frequencies, it can be taught to avoid transmitting other undesirable

disturbances as well. The grip is tightened for joint maneuvers, but relaxed between them, so that payload and bus actually fly along identical orbits without touching. More than just a precise positioning device then, the space bearing is a mechanical switch, able either to isolate completely our payload mirrors or transmit motions to them, as the fleet controller commands.

The selection of actuator technologies reviewed above clearly provides sufficient variety of method, speed, strength, stroke and precision to establish the feasibility of relative, controlled movement among any portions of our satellite structures, and the mirrors they carry. As precedent for our hierarchical actuator scheme (first rigid-body, then thermal, PZ, and finally EM) we note in fact some demonstrated active members developed specifically for COSS, which by combining different actuators in one unit can adjust relative position with both coarse and fine tuning [Fanson & Chen, 87] [Hoehne, 84]. Controlling the shape, pointing and vibration of LSS which have "little or no intrinsic out-of-plane stiffness" is therefore simultaneously possible with distributed active structural elements [Chen & Fanson, 85].

Sensors

"The use of distributed actuators leads quite naturally to distributed sensing" [Szirmay, 79]; in fact, accuracy is increased and processing complexity reduced if sensors and actuators are collocated to give a system the best possible chance of knowing exactly what it's doing [Baruh, 85]. Very few sensor types, however, are so compatible with actuators as to allow rigorous collocation, so we do not simplistically require a number of sensors equal to the actuator number. Our ultimate sensing purpose is to measure the "three fundamental" properties of the millions of mirror elements defining our resonator: spacing (line of sight range), alignment (transverse displacement), and collimation (relative tilt) [Yager, 78], thus allowing their consequent actuation. In pursuit of these data, we must employ a variety of measurements.

Beginning with sensors operating in close proximity to their ultimate measurands, we distinguish for our present vacuum robotic application the non-contact, or field, sensors (magnetic and electric) from contact, or point, sensors (force, torque, pressure, position, temperature) [DePaula et al, 87]. We focus at once on fiberoptical sensors, because in general they represent the most advanced (versatile, accurate and economical) and rugged of sensor types, with the tremendous additional advantage that they are directly integrable into fiberoptic data transmission and processing systems.

Although a plethora of non-interferometric properties can serve as the basis for fiber sensors (attenuation and scattering, to name just two), to the end of establishing feasibility for our uses, we consider here the most sensitive methods, and in particular the most developed, Mach-Zehnder interferometry. Ultimately the measurand can be any physical property capable of effecting geometrical change in some

material. If that material comprises a mandrel wrapped by the measuring fiber, or if it clads the fiber directly, its response to the changing measurand will alter the fiber strain or refractive index. An axial strain or index difference between affected and control fibers can then be measured by direct phase interferometry; presently hybrid Integrated Optical Circuits (IOCs) perform this fringe analysis, but even lower-loss all-optical componentry is expected for the future. Radial strain changes, which affect fiber birefringence, are measured independently but simultaneously.

Traditional electrical strain sensors (whose resistances change with Poisson distortions when strained) cannot approach within orders of magnitude the sensitivity of such optical sensors. Using light of wavelength $1\text{ }\mu\text{m}$ (near-IR is commonly used in fiberoptic systems because both laser diodes and low-loss fibers exist for this range) "one can detect phase shifts equivalent to a displacement of 10^{-13} meters" [DePaula et al, 87]. Thus a "squeezed fiber need only generate a relative Poisson elongation on the order of a nuclear dimension to be detected with a fiber optic interferometer." Although techniques used for gravity wave research can surpass even such amazing performance, it would be hard to imagine a more rugged and unobtrusive sensor than some pieces of optical fiber attached to an IOC. It makes good sense that metrication of an optical robot would be best effected using optical means.

Fiberoptic magnetic sensors can resolve field strengths of 50 pT already, within three orders of magnitude of their theoretical best performance. Such sensitivity clearly offers possibilities both for ambient field sensing and for non-contacting kinematic sensing. All-optical rotation sensors which can resolve rates down to $5(10^{-3})\text{ deg/hr}$ (still also three orders of magnitude away from their theoretical limit) [DePaula et al, 87] provide other useful kinematic measurement

possibilities. Another advanced kind of interferometric, non-contacting rotation sensor not based on fiberoptics uses small, stabilized two-frequency HeNe laser readout heads to resolve rotations of a grating-encoded cylinder as small as a few tens of nrad, at 225 kHz update rates, for rotations as fast as 5 rad/s [Hercher & Wyntjes, 87]. These techniques give us ways of measuring the relative displacement of adjacent mirror segments.

Accurate intra-structure displacement measurement is impossible without simultaneous detailed temperature knowledge, since the dimensions of all materials change somewhat with temperature variations. "To make a sensor which is sensitive to temperature is probably one of the easiest activities in which one can become engaged. However, to make a good sensor for measurands other than temperature without having that sensor influenced...by temperature is extremely difficult" [DePaula et al, 87], as is the reverse problem. Right away this tells us that our system must use different types of fiberoptic temperature sensors (which are legion) to derive useful temperature information, and their calibration must be continually verified as environmental degradation changes their materials. Current resolution is 10^{-4} K (four orders of magnitude away from the theoretical best), good to over 2200 K.

Although the range of phenomena measurable by fiberoptic sensors is seemingly limited only by ingenuity (they have been developed also for radiation, current, and chemistry), it is of course in the direct measurement of structural strain that they excel. Coincidentally, strain is recommended as a much more reasonable measurand than small velocities for LSS [Fanson & Chen, 87]. With yet greater freedom than conventional strain gauges, fiberoptic strain sensors can be bonded to virtually any structure, and even embedded directly in composite layups

[Rowe, 86]. Their rugged ability to be incorporated into many materials (even into PZ ceramics themselves) means that not even the piezoelectric actuator-sensors discussed already are more collocatable than a well-engineered fiberoptic sensor. With embedded and applied strain fibers, then, we can derive continuum deformations throughout the fleet structures.

Inertial sensors (accelerometers) yield absolute kinematic data, rather than the relative data provided by optical sensors [Aubrun et al, 83]. Rate Gyroscope Assemblies (RGAs) based on angular momentum conservation and conventionally used for spacecraft attitude control are prey to both mechanical and electrical noise; refinements have reduced error in the Space Telescope gyros to about 0.1 deg/hr [Dougherty et al, 83]. But an accelerometer developed at the University of Maryland for NASA represents the new approach [NASA TB, 8707/8b]. It detects simultaneously the 6-DOF motions of a magnetically-suspended precision cube proof-mass with a single amplifying SQUID (Superconducting Quantum Interference Device), resolving linear accelerations as small as $pggHz^{-\frac{1}{2}}$, or three to six orders of magnitude better than conventional instruments (where $g_E \equiv 9.8 \text{ m/s}^2$). Angular acceleration sensitivity is of the order $nrad \cdot s^{-2}Hz^{-\frac{1}{2}}$, with short term stability exceeding gyros. The instrument does, however, still require a cryogenic environment.

The Department of Defense knows of several promising inertial technologies and even proven devices which are almost as amazing and much cheaper, the best of which use the proof-mass design [Aubrun et al, 83]. Two in particular, developed by Charles Stark Draper Laboratories, are called SASS (Six Axis Space Sensor) and TAARA (Three Axis Angular Rate and Acceleration sensor). Both come in packages just centimeters across, optimized for microgravity use, rugged and costing much less than gyros; "compatible with optical sensing techniques",

they deliver linear position resolutions down to a few nm and angular position resolutions down to tens of nrad, from DC (linear) and 0.1 Hz (angular) to 100 Hz. Linear acceleration sensitivity is tens of ng_E (hundreds of nm/s^2), and angular acceleration sensitivity is a few $\mu rad/s^2$. Locating such small inertial sensors at the mirror attachment points will allow another cross-check of force disturbance.

At first it would seem that accelerometers could measure only local disturbances (such as a propagating wave due to some mechanical forcing, which passes through the structure), and not global field perturbations (such as overall satellite motion resulting from gravity field bumpiness). After all, the force of gravity varies with an object's mass, so a changing planetary tug will accelerate both the bus and an accelerometer's proof mass identically, yielding no data on the field (or the spacecraft's position relative to the planet). However, a device derived from the SQUID accelerometer discussed above [NASA TB, 8711/12], by using several proof masses in a differential mode, can measure gravity gradients directly along three separate axes. It reads out three common-mode linear accelerations and three in-line component gradients simultaneously, combinations of which are used to obtain "precise position" information, "valuable cross-checks of the gravity data", and a cross-check of other gyro attitude rate data, at a sensitivity up to five orders of magnitude better than competing gradiometers. For us this means that the spacecraft can know their positions relative to the radial and tangential gravity variations of Venus. Locating such devices across the extent of the large satellites will enable even greater coupled accuracy as the craft fly along the contours of the planetary gravity field.

Thus technology to characterize through "proximal" sensing the mechanical behavior of adjacent, self-contained mirror

elements and their supporting structure both exists and continues to improve rapidly. But our system features expanses of such hardware kilometers in extent, themselves comprising islands separated by thousands of kilometers of space, which must still work together. Next we must look into "distal" sensing methods for mirror control. Since these are by definition non-contacting, they must use radiation to transfer information and hence are optical.

Laser ranging techniques are well-developed for geodetic and tactical military applications. The LAGEOS satellite orbits at about 6000 km altitude, a range of the same order as our 9000 km satellite separation. Demonstrated single-shot precision is 0.9 m, "close to the theoretical value" [Paunonen, 82]. Military emphasis has been on small, low-power, rugged and lightweight technology; typically a 2 kg instrument the size of a pair of binoculars using a Nd:YAG laser can range to 10 km with a resolution down to 3 m [Johnson, 86] [Daly, 86]. Since this type of ranging calculates distance from photon time-of-flight, its accuracy is limited by the laser pulse shape and clock frequency. Modest improvements are expected, but mostly in reliability and portability.

An intermediate capability, for closer range but with greater accuracy, measures lateral deviations of bright, modulated point sources with planar photodetectors at the foci of dedicated tracking telescopes. In one demonstration such a sensor tracked six targets simultaneously at a range of 4.5 m with an accuracy of 10 μ m, using all off-the-shelf hardware [Neiswander, 79]. This system, designed in fact to provide figure error monitoring for segmented space-based reflectors, is subject to refinements improving its performance by at least a factor of ten [Neiswander, 78]. A similar centralized scheme is being developed for the 20 m far-IR Large Deployable

Reflector (LDR); a fiberoptically enhanced time-of-flight method yields range uncertainties of $150\text{ }\mu\text{m}$ and angular uncertainties of $100\text{ }\mu\text{rad}$ over distances of about 20 m [Dahlgren & Taylor, 84]. These ranging techniques can enable control neighborhoods across the face of each of our resonator satellites to monitor each others' relative positions.

The final category of optical metrology is again based on interferometry and yields the best possible results. Coherent light sent along two paths is interfered; analyzing the resulting fringe motion accurately gives a direct measure of the paths' relative length change. Commercially available HeNe laser interferometers can achieve accuracies of about one part in 10^7 for relative speeds up to 18 m/min [Berkman, 79]. This "continuous length measurement over distances of many meters to submicrometer accuracy" is made possible with frequency-stabilized laser sources having coherence lengths over 1 km [Malacara, 78]. Heterodyne techniques have been used to expand the resolution dynamic range. For instance, a DoD scheme used a two-color CO_2 laser to generate a hierarchy of long "synthetic" wavelengths capable of measuring, at several Hz, absolute separation over kilometers with sub- μm accuracy by "handing off" the fringe count from coarser to finer measurement ranges [Davis et al, 78]. And Mottier [78] reported an argon laser system developed to scan 21 target points (on an adaptive mirror for a CO_2 laser) every 120 msec, with electronic phase interpolation of interference fringes yielding 10 nm resolution.

Hewitt [84] points out several important aspects of interferometric metrology. First, the uncertainty of a length measurement is generally an order of magnitude worse than the resolution of a displacement measurement; but for our most stringent purposes, relative displacement is what we need to measure anyway. Second, the daunting task of aligning

perfectly a retroreflector so as to return the mensuration beam to the interferometer can be avoided completely by using passive corner cube reflectors, which of course automatically retroreflect. Presuming, then, displacement metrology from some calibrated condition, even the cube mounting need not be precise, as long as it does not change. Such interferometric methods allow the fleet controller to monitor displacements among entire regions of control neighborhoods across the resonator craft.

While a combination of the techniques outlined above can clearly meet our needs for monitoring the relative positions of mirrors across the vast surface of each of our resonator satellites, even the enhanced interferometric resolution of about one part in 10^9 falls hopelessly short of the 1 in 10^{15} resolution needed to track 62 nm excursions in a 45(10³) km resonator path. And the best-behaved satellites imaginable cannot effect a planetary laser unless they work cooperatively as one machine. Now, it is the finite coherence length of radiation which limits interferometric resolution, which is why the field did not really develop until lasers appeared [Hewitt, 84]. But the active robot which our laser system is, necessarily has available (when operating) a laser with the longest coherence length in history, namely the cavity beam itself. We can use this resource as follows: the circulating field contains at all times sufficient phase information to "know" whether or not it remains in phase with itself --- otherwise it would not constitute a laser. Thus a resonator station employing phase-sensitive interferometric detection would be able to use that coherent source to monitor its displacement relative to the beam it helps define. A voting array of stabilized interferometric sensors distributed across each satellite will therefore enable its segmented mirror surface to "ride" the resonant cavity wave, maintaining the beam's coherence.

Although they pertain more aptly to bus attitude control, we list here the other sensor types required for fleet operation. Tuned infrared planetary "limb" sensors allow the controller's master plan to keep the cavity beam conservatively centered within Venus' mesospheric inversion layer. Dedicated inter-craft laser telemetry links double for aiming purposes, their signals collected by small (20 cm) telescopes [LaPrade, 87]. Simple photodiode overspill sensors collaring the Transducers, Rings, and the 1δ and 1ε craft enable them to remain actively centered on the actual output beam and communicate repointing information throughout the fleet. The ultimate attitude reference for the entire fleet is of course distant stars. Several fixed-head, voting, telescopic star trackers mounted on the payload sides of the craft's final EM isolation stages decide on and update a reference fleet attitude continually. The precision of such sensors is a function of prior astrometric accuracy (the star "charts" in their memory); expected to attain 485 nrad even by 1991 [Laskin & Sirlin, 86], the star trackers themselves should not limit pointing accuracy for a planetary laser. Sun sensors, being far less accurate, are not used in our fleet.

A combination of inertial and optical sensors sprinkled liberally around the fleet thus provides the fleet controller with all the state intelligence it needs (at all levels of detail) to operate the planetary laser continuously for its changing missions over many years.

Appendix A8-1 Orbital perturbations due to non-spherical Venus.

The most comprehensive study of Venus' gravity field yet done is the one by Bowin [85] referenced in Appendix A6-6. That study states explicitly that its model cannot resolve gravitational features with a half-wavelength smaller than 5° (of planetary longitude or latitude). Bowin develops contour maps for the planet showing constant-altitude gravity-induced radial and tangential accelerations, and identifies for exclusion those features which are spurious artifacts of the data processing. The most severe anomalies estimated from those maps for an equatorial orbit track are $\approx 20 \text{ mgal}/15^\circ$ at 300 km altitude for tangential acceleration, and $\approx 40 \text{ mgal}/15^\circ$ at 500 km altitude for radial acceleration. $1 \text{ mgal} \equiv 1 \text{ cm/s}^2$ is the traditional unit of acceleration used in planetological gravity study. Following the scale-factor method developed in Appendix A6-6, we derive attenuated acceleration values for our orbital altitude:

$$\left(\frac{300 + 6052}{7641} \right)^{1/4} (20) = 1.5 \text{ mgal}/15^\circ \text{ (tangential)}$$

$$\left(\frac{500 + 6052}{7641} \right)^{1/4} (40) = 4.6 \text{ mgal}/15^\circ \text{ (radial)}$$

For a circular orbit, orbital anomaly translates directly as time, so we are really representing jerk (acceleration per unit time) here. 15° of our orbit take $7363/24 = 307 \text{ s} \approx 5 \text{ min}$. Assuming most simplistically (and conservatively) that the jerks derived above are constant, we can easily calculate the maximum non-Keplerian excursions that would result during just that interval.

$$d_{\max} = \frac{1}{2} a t^2 = \frac{1}{2} (.015) (307)^2 \approx 710 \text{ m (tangential)}$$

$$= \frac{1}{2} (.046) (307)^2 \approx 2200 \text{ m (radial)}$$

The latter result confirms our conclusion of Appendix A6-6, which was based on a different type of anomaly map in Bowin's report.

Realistically, because of the nonzero rise-times of anomalous gravitational events, we should expect the actual non-Keplerian excursions to be rather smaller than these values.

Appendix A8-2 Gravity gradient torque at Station 1.

Gravity gradient torques on orbiting spacecraft come in three types: secular, cyclical, and constant. Secular (accumulating) and cyclical (cancelling) torques result if the craft flies inertially stabilized, essentially ignoring its orientation with respect to the planet it orbits. The satellite and its primary can be thought of as a gravitational machine; to avoid the effect of the gravity gradient (which would normally planet-orient the satellite), the craft must be "wound up" rotationally by an external torque. Secular terms in the torque vector can only be compensated by continual addition of energy, generally either propulsive or magnetic. Cyclical terms neutralize themselves over the orbital period, and can in general be compensated temporarily by storing angular momentum onboard.

If the craft flies planet-oriented, the gravity-gradient "engine" is in neutral and no varying torques develop. If in fact it flies principal-axis-oriented (inertially symmetrical with respect to the orthogonal orbit coordinates), then no gravity torques occur. The satellite attitude is either in conditional or stable equilibrium (the latter if the moment of inertia around the axis lying along the orbit radius is smaller than the others). Our basic vertex stations fly in conditional equilibrium with respect to the gravity gradient.

But the large satellites of Station 1 all fly non-principal-axis-oriented. Although they are planet-oriented in the sense that they rotate once per orbit, their principal inertial axes are rotated with respect to orbital coordinates. The gravity-gradient engine is still in neutral, but at a biased, or constant, torque level. The crafts' attitudes, required by optical considerations, continuously fight the gravity gradient's attempt to align their principal axes with the orbital coordinate axes. To evaluate ways of supplying the necessary external torque to maintain their attitudes, we must estimate its size.

The gravity gradient torque increases with spacecraft mass, dimension, and departure from a principal-axis orientation. 1β is at once the largest of the Station 1 spacecraft and the most extremely tilted with respect to the orbital coordinate frame. Thus we use 1β to estimate the size of the effect.

Chapter 4 does not develop in detail a reference design for 1β, since its systems are much the same as those of Stations 2, 3, 4 and 5. Thus we loosely adapt design data from those craft to estimate 1β's inertial properties. 1β's reflector ellipse is 1000 x 1140 m, an area only 2/3 as large as those others'. We use this factor to reduce their mass value, since the overall satellite mass for these reflector craft varies closely with their area. Furthermore, we estimate that the overall radius is about 650 m, compared to their 850 m. Moment of inertia is proportional to mass and to the square of dimension, so we use the factor:

$$(.67) \left(\frac{650}{850} \right)^2$$

to adjust the moments of inertia calculated in Appendix A4-1 for our application to 1β. We also make I_{zz} more closely match I_{yy} to reflect 1β's virtual circularity:

$$I_{xx} \approx 12 \text{ T kg m}^2$$

$$I_{yy} \approx 5 \text{ T kg m}^2$$

$$I_{zz} \approx 6 \text{ T kg m}^2$$

These moment of inertia terms represent the diagonal entries in the matrix representation of the spacecraft inertia tensor, as evaluated for the principal axis configuration. The off-diagonal, or product of inertia, entries are all zero in this symmetrical case.

The gravity gradient torque vector T is given by [Woodcock, 86]:

$$T = \frac{3\mu}{r^3} r \times [I] \cdot r \quad (A8-2.1)$$

where the vector r is the normalized, dimensionless orbital radius, measured of course in the same frame as the second-rank inertia tensor $[I]$. In orbital coordinates (which move with the satellite):

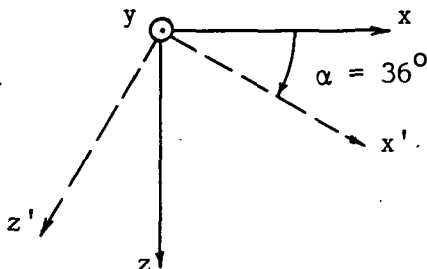
$$r = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

for planet-oriented motion. To get $[I]$ into the same coordinate system, we must employ a coordinate transformation:

$$[I'] = [A][I][A]^T \quad (A8-2.2)$$

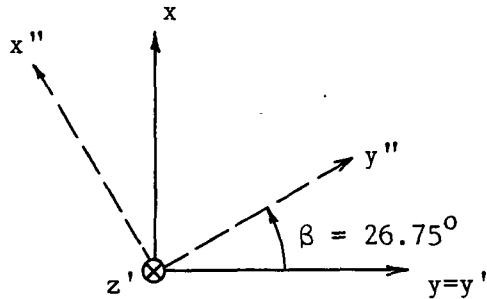
where the transformation matrix $[A]$ must be assembled [Woodcock, 86] from matrices encoding each successive Euler angle rotation of the spacecraft away from its principal-axis orientation.

The convention for orbital coordinates is to have the x axis along the orbit track, the y axis cross-track, and the z axis nadir-pointing. This convention coincides exactly with our typical choice of bus coordinate axes if we begin with β 's reflector normal along the orbit path and its ellipse major axis cross-track. Thus our nomenclature for the inertia terms is already proper. First we rotate the satellite around the y axis (pitch it down) 36° :



$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Next we rotate it around the z' axis (negative yaw in this coordinate convention since we chose the original x axis to be antiparallel with the orbit velocity vector) 26.75° :



$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

This second rotation matrix must premultiply the first, according to the rules of linear algebra, because it acts on the result of that one. Performing that operation yields the transformation matrix:

$$[A] = \begin{pmatrix} \cos\beta\cos\alpha & -\sin\beta & \cos\beta\sin\alpha \\ \sin\beta\cos\alpha & \cos\beta & \sin\beta\sin\alpha \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix}$$

The original, principal-axis inertia tensor is:

$$[I] = \begin{pmatrix} 12 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} (10^{12}) \text{ kg m}^2$$

Substituting both $[A]$ and $[I]$ into equation A8-2.2 produces a symmetrical matrix representing the non-principal-axis inertia tensor $[I']$. A full matrix, it now includes product-of-inertia terms:

$$I'_{11} = 12\cos^2\alpha\cos^2\beta + 5\sin^2\beta + 6\sin^2\alpha\cos^2\beta$$

$$I'_{12} = I'_{21} = 12\cos^2\alpha\sin\beta\cos\beta - 5\sin\beta\cos\beta + 6\sin^2\alpha\sin\beta\cos\beta$$

$$I'_{13} = I'_{31} = -6\sin\alpha\cos\alpha\cos\beta$$

$$I'_{22} = 12 \cos^2\alpha \sin^2\beta + 5\cos^2\beta + 6\sin^2\alpha \sin^2\beta$$

$$I'_{23} = I'_{32} = -6\sin\alpha\cos\alpha\sin\beta$$

$$I'_{33} = 12 \sin^2\alpha + 6\cos^2\alpha$$

all of which terms are multiplied by the common factor and units $(10^{12}) \text{ kg m}^2$ which we have left outside the matrix. Substituting the angles 36° and 26.75° for α and β respectively, we can fill out $[I']$ numerically:

$$[I'] = \begin{bmatrix} 8.9289 & 1.9803 & -2.5478 \\ 1.9803 & 5.9982 & -1.2842 \\ -2.5478 & -1.2842 & 8.0729 \end{bmatrix} (10^{12}) \text{ kg m}^2$$

We now substitute this transformed inertia tensor, the normalized radius vector, and numerical values for our orbital radius r (7641 km) and Venusian gravitational parameter μ_\oplus ($324.86(10^{12}) \text{ m}^3/\text{s}^2$) into equation A8-2.1 to calculate the gravity gradient torque vector, after matching units:

$$\mathbf{T} = \begin{bmatrix} 2.81 \\ -5.57 \\ 0 \end{bmatrix} \text{ MN m}$$

This shows the constant torques we may expect in orbital coordinates, acting on 1β . As we would expect, the torque about the nadir axis is zero; because the first-order gravity gradient is radial only, it cannot twist the satellite's attitude about the radius vector.

Compensating the Torque

The other torque components are quite large, however. Assume first that they are to be provided propulsively. Given a moment arm equal to the 650 m bus radius, a non-translational couple for the 5.57 MNm torque alone would require two engine stations, each delivering 4282 N. If provided by our efficient ion engines, this force would call for 3900 1.1 N engines at each station, consuming a constant 234 MW of electrical power, and 6,125 MT of propellant each year. The propulsive solution is one we should avoid.

Magnetic torquing at Venus would be ineffectual for such a large torque. First of all, the weak field varies greatly due to solar wind interactions. Second, the available torque is $T = nIA \times B$, where $n \equiv$ number of current loops, $I \equiv$ current applied, $A \equiv$ enclosed area vector given the current sense, and $B \equiv$ ambient magnetic induction. A flat satellite orbiting planet-oriented cannot achieve 3-axis control from the planetary magnetic field. Finally, even given the large area available on 1 β , the weak planetary field and large required torque mandate a minimum of $2.22(10^6)$ ampturns, quite a systems penalty.

The clever, and in fact only reasonable, option is structural. If the satellite is part of a tethered system (where the "tether" could in fact be a long rigid structure, if desirable for other reasons), a sufficiently large tether tension, located sufficiently far away from the bus CM, can provide a constant attitude "disturbance" torque for as long as the configuration orbits. In essence, this approach uses the gravity-gradient engine which an orbital tether system is to counteract the effect of the gravity-gradient engine created by a non-principal-axis satellite orientation. Both engines are in neutral, both run at a biased constant torque level, and they cancel each other out. It makes sense that only a passive mechanism tapping the vast gravity field of the planet could compensate a disturbance arising from that same source. Such structural control of gravity-gradient torques is used throughout Station 1.

Appendix A8-3 Third-body perturbations.

The third-body gravitational effects of the sun are not negligible even for small satellites orbiting the Earth at altitudes above about 700 km. The sun's effective gravitational potential U_{eff} is found [Kaplan, 76] as:

$$U_{\text{eff}} \approx \frac{\mu_0 r^2}{2r_0^3} (3 \cos^2 \phi - 1) \quad (\text{A8-3.1})$$

where $r \equiv$ satellite orbit radius, $r_0 \equiv$ satellite primary's orbital radius about the sun, and $\phi \equiv$ instantaneous angle between those two vectors as the three bodies move. The effective potential is obviously maximized when $\phi = 0$ or π , when the three bodies are collinear, so that $3 \cos^2 \phi - 1 = 2$. Then:

$$U_{\text{eff,max}} = \frac{\mu_0 r^2}{r_0^3}$$

The effective force per unit satellite mass (its acceleration) is found by differentiating:

$$a_{\text{eff,max}} = \frac{\partial}{\partial r} (U_{\text{eff,max}}) = \frac{2\mu_0 r}{r_0^3}$$

Substituting relevant values yields the maximum instantaneous perturbing acceleration:

$$a_{\text{max}} = \frac{(2) \cdot 1.327(10^{11}) \cdot 7641 (10^3)}{(108.1(10^6))^3} = 1.6 \frac{\mu\text{N}}{\text{kg}} = 1.6 \frac{\mu\text{m}}{\text{s}^2}$$

The solar tug which is along the orbit track (near the terminator crossings) cancels with each orbit, since energy which is added by accelerating the satellite tangentially during one half of the orbit is then subtracted by decelerating the satellite tangentially during the other half. The radial acceleration, however, is secular since it acts during both dayside and darkside passages to pull the orbit out-of-circular, changing not its energy (size) but rather its eccentricity (shape). We can consider this radial perturbation as the absolute value of a sinusoidal function, varying from zero up to a maximum of $1.6 \mu\text{N/kg}$ as we just calculated, twice per orbit. We seek an average value for the force, to simplify propulsion logistics analysis. The area under one "hump" of a standard sine curve is 2 , so the height of the rectangle with length 2π which has equivalent area to two humps is $4/2\pi = 0.64$. Thus the constant-force equivalent for the maximum tug we have calculated is $(.64)(1.6) = 1.0 \mu\text{N/kg}$.

Local Bodies

For the component satellites of Station 1, a much larger source of third-body gravitation than the sun is each other. That is, 1α , 1β and 1γ are huge craft containing much mass, which orbit in close formation with each other and with the tethered pair of 1δ and 1ϵ . The usual practice of ignoring satellite mass would for us be a hazardous oversimplification. For example, referring back to equation A8-3.1, we can estimate the effective gravitational potential 1 km away from 1γ , assuming $\phi = 90^\circ$ (along the orbit path):

$$a_{\text{eff}} = \frac{\mu_\gamma r}{r_\gamma^3} = \frac{6.67(10^{-11})(.75)95(10^6)7641(10^3)}{1000^3} = 37 \frac{\mu\text{N}}{\text{kg}}$$

where we have assumed 1γ 's mass to be 75 % of one of the basic vertex stations (Appendix A4-1), reasonable since the system mass for the large flat resonator craft scales with their area. This gravitational

acceleration is 37 times larger than that due to the sun, and thus dominates third-body effects at Station 1.

We choose two approaches to meeting the problem. Our reference design for the tethered pair counteracts this force propulsively. To first order, the radial effects on 1δ and 1ϵ can be ignored, even though they ride below and above the resonator orbit, since their combined CM rides on the orbit. To first order we will also ignore the tangential effects due to 1α and 1β , since they are about the same size and symmetrically placed about the tethered pair. But 1γ is an uncancelled source of gravity, which produces approximately the acceleration calculated above; the system repercussions of dealing with it propulsively are addressed by Appendix A4-4.

Propulsion is an inefficient way of handling the problem for 1α , 1β and 1γ themselves, however, due to their large masses. That is, each feels a total force on the order of $37(10^{-6})(.75)95(10^6) = 2600 \text{ N}$ due only to the next closest large satellite. Meeting this propulsively would require thousands of ion engines expending tonnes of propellant each day. The slight mass penalty of some extra (even active) structure to connect these three craft and keep them apart is much smaller than the enormous power, mass and complexity penalties introduced by thousands of engines blasting xenon ions at each other and requiring staggering logistical resupply. Therefore we will refer to our solution of the inter-gravitational formation-keeping problem as a "structural" one.

Appendix A8-4 Non-Keplerian orbit penalty for 1γ .

The CM of the 1γ bus is displaced from the nominal Keplerian resonator orbit in two ways: it travels about 250 m below the orbit, and about 1000 m south of the orbit plane, as noted in Chapter 7. Unattended, these desired displacements would not remain fixed. Rather, to first order, a satellite stationed below another would move faster, building up an alarming along-track separation in only days. And a satellite stationed out-of-plane would travel in a planet-centered orbit tilted relative to the nominal orbit, so that the path separation oscillated southward and northward of the orbit plane with each revolution. The formation requirement of the planetary laser fleet can admit neither of these "natural" departures. 1γ 's relative position must be maintained artificially.

Altitude Penalty

We estimate the penalty using an energy approach. The specific mechanical energy of a circular orbit is given by:

$$E = - \frac{\mu}{2a}$$

so that the energy difference (which must be made up from an external source) between two closely-spaced orbits is approximately:

$$\Delta E = \frac{\mu \delta}{2a^2} \tag{A8-4.1}$$

where $\delta \equiv$ the altitude difference between the orbits, and $a \equiv$ the nominal orbital radius (semi-major axis). The specific mechanical power necessary over the orbital period T to supply this energy is:

$$P = \frac{\mu \delta}{2a^2 T} \quad (\text{A8-4.2})$$

The thrust T provided by electric propulsion is given as:

$$T = \frac{2 \eta P}{I_{sp} g_0} \quad (\text{A8-4.3})$$

where the product $\eta P \equiv$ the input electrical power reduced by the system conversion efficiency (typically about 0.8 for ion engines), $I_{sp} \equiv$ the specific impulse, and $g_0 \equiv 9.8 \text{ m/s}^2$. Substituting equation A8-4.2 for ηP in A8-4.3 tells us the specific force averaged over an orbit necessary to displace 1γ 250 m below the nominal orbit:

$$T = \frac{(324.86)(10^{12})(250)}{(7641(10^3))^2(7363)(4500)(9.8)} = 4.28 \text{ } \mu\text{N/kg}$$

Assuming that 1γ is about $2/3$ the mass of a basic vertex station (see Appendix A8-2), the total average force required for altitude displacement is:

$$4.28(10^{-6})(.67)95(10^6) = 407 \text{ N}$$

which could be managed at any given time by 370 1.1 N ion thrusters operating at maximum power (a total of 11.1 MW) and consuming 291 MT of propellant each year.

Alternatively, 1γ could be maintained at its proper altitude passively by tethering it to a counter-mass stationed appropriately above the resonator orbit, as with the $1\delta - 1\epsilon$ pair. The same tether could, if attached at the proper point, provide the torque necessary to cancel 1γ 's gravity gradient torque, as explained in Appendix A8-2. Indeed we choose this structural approach rather than the propulsive one.

Out-of-Plane Penalty

We follow exactly the procedure outlined in Appendix A5-2. Letting M be the total mass of 1γ , the separation force required to displace it 1000 m south of the orbit plane at all times is:

$$F_s = \frac{\mu_\oplus M}{a^2} \sin\gamma = \frac{(324.86)(10^{12})(.67)95(10^6)}{(7641(10^3))^2} \frac{1000}{7641(10^3)}$$
$$= 46.3 \text{ kN}$$

which is quite large and could not be handled by any reasonable mass expulsion technique we know of.

We must buy our way out of this force penalty with, again, structure. We cannot simply "tether" a counter-mass on the other side of the orbit plane because the linkage must carry compressive force, rather than the tensile force characteristic of a gravity-gradient stabilized arrangement. Still, the mass penalty of a counter-satellite and the structure to pair it with 1γ , whether that structure is actively or quasi-actively stiff, must be much smaller than the mass, complexity, expense, reliability and logistics penalties associated with the propulsive alternative.

Appendix A8-5 Drag perturbation due to Venusian exosphere.

Data on atmospheric density at great heights above Venus are not readily available. Wertz [84], however, provides a logarithmic plot for Earth's atmospheric density which is easily extended to our orbital altitude of 1589 km. For comparison, that source indicates a density of about 10^{-9} kg/m³ at 170 km; since Ananda et al [80] show data indicating a density one order of magnitude less than this at the same altitude above Venus (subsolar, which should already be the greatest because of dayside swelling), using the Earth data will yield a conservatively excessive drag value. We project 10^{-17} kg/m³ at our resonator orbital altitude. The drag force per unit frontal area encountered while moving through this density is given by:

$$F = \frac{1}{2} \rho v^2 C_D = (10^{-17})(6520) = 4.25(10^{-10}) \text{ N/m}^2$$

where we have assumed the maximum value of 2 for the drag coefficient C_D , and substituted our orbital velocity v . The planet-facing basic vertex stations present comparatively little frontal area to atmospheric drag; satellites 1 α , 1 β , and 1 γ suffer the most. Taking 1 β as an example, we evaluate its projected frontal area and calculate the upper bound drag force to be:

$$4.25(10^{-10}) \pi (500)(550) \cos 36^\circ \cos 26.75^\circ \approx 3(10^{-4}) \text{ N}$$

This value represents the maximum continuous thrust which station-keeping propulsion must be able to provide to neutralize the drag.

Appendix A8-6 Radiation pressure effects throughout the fleet.

Intermediate Beam

The 180 kW output beam, when focused down to a 10 m diameter, has power density:

$$\frac{180(10^3)}{\pi(5^2)} = 2300 \text{ W/m}^2$$

which is somewhat less than the 2660 W/m² sunlight strength at Venus but rather larger than the 1390 W/m² sunlight strength at Earth.

Reflected Sunlight

Venus' Bond albedo, defined as the ratio of total light reflected to total light intercepted, is 0.72 [Wertz, 84]. Because the planet intercepts a disk but radiates as a hemisphere, its average reflected radiance will be only half the local subsolar maximum brightness. But that maximum value is of interest for spacecraft engineering, and its effective value at our orbital altitude is:

$$\left(\frac{6052 + 58}{7641} \right)^2 (.72) 2660 = 1225 \text{ W/m}^2$$

where we have taken the radiating surface as the top of Venus' reflective H₂SO₄ cloud layer.

Thermal Emission

Venus' effective temperature, fairly constant over the entire planet, is 231 ± 7 K [Hunten et al, 83], so using the Stefan-Boltzmann law

we find a spacecraft at our altitude subjected to a flux of:

$$\left(\frac{6052 + 58}{7641}\right)^2 5.67(10^{-8})(231)^4 = 103 \text{ W/m}^2$$

Combined Effect

The summed radiation pressure on our resonator satellites is less severe than the sunlight pressure alone would be, because of geometry. We will examine two "worst" cases. The solar pressure dominates, and exerts the largest integrated force on the basic vertex stations, which have the largest area. But their area is normal to sunlight as they pass the subsolar point, when both the planetary thermal emission and maximum reflected radiance oppose the solar force. The total is:

$$\frac{\pi(500)(850)}{3(10^8)} \left\{ (1.7)2660 - (1.83)1225 - (2)103 \right\} = 9.2 \text{ N}$$

where a solar reflectance of 0.7 is assumed for the rear face of the spacecraft (Appendix A9-2), a (reflected) solar reflectance of 0.83 is assumed for the gold-fronted mirrors [Berggren & Lenertz, 75], and the thermal (infrared) reflectance of the mirrors is of course practically unity.

Alternatively, when Station 1 overflies the terminator, reflected radiance fades, and planetary thermal emission no longer mitigates the solar pressure since they do not align. But both exert substantial forces since the 1 β and 1 γ satellites are never edge-on as the basic vertex stations can be. Specifically, 1 β experiences its greatest solar pressure when it is 36° (equal to its pitch bias) past the terminator plane, at which point it is still in full sunlight (Appendix A8-6). Furthermore, in that position both the solar and planetary fluxes strike its highly reflective gold front, maximizing momentum

transfer. Found using vector algebra, the combined force is:

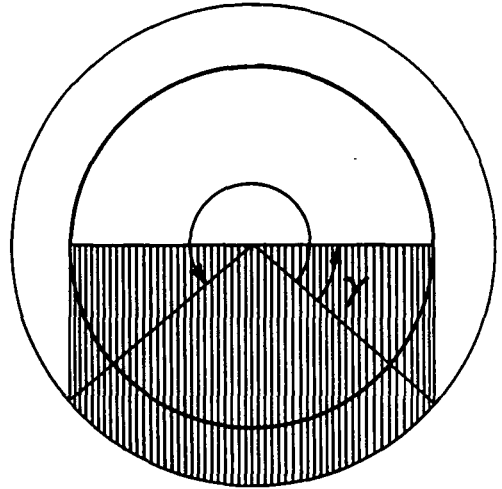
$$\frac{\pi(500)(550)}{3(10^8)} \cos 26.75^\circ \left\{ ((1.83)2660 + (2)103 \sin 36^\circ)^2 + \right. \\ \left. + ((2)103 \cos 36^\circ)^2 \right\}^{\frac{1}{2}} = 12.8 \text{ N}$$

To effect simultaneous compensation, the resonator satellites must carry stationkeeping propulsion systems capable of meeting these respective maximum thrust values.

The mirror areas of the 1δ and 1ε craft are so small by comparison with the rest of the resonator satellites that the total radiation pressure they experience, from the sun and the intermediate beam, is of order mN. The Transducers and Rings experience even less, since diffraction spreading weakens the intermediate beam by 4/5 before they intercept it (Appendix A7-16). Planetary sources become truly insignificant at the libration points, and even the sun is partially obscured by Venus for L2 Station.

Appendix A8-7 Resonator orbit eclipse and terminator impulse.

Because of the orbit's altitude, much more than 50 % of it is in sunlight. The top of Venus' opaque H_2SO_4 cloud layer is taken as 58 km above the planetary surface. Combining this with our orbital geometry yields the angle γ :



$$\gamma = \cos^{-1} \left(\frac{6052 + 58}{7641} \right) = 37^\circ$$

Then $\frac{2\gamma + 180}{360} = 71\%$ of the orbit is sunlit.

That interval corresponds to	5191	out of	7363 s
	86.5	out of	123 min
	1.44	out of	2.05 hr

Given our reference fleet formation, the 1 α satellite (with a pitch bias of -36°) enters sunlight with a practically normal (within 1°) orientation, its highly reflective gold surface facing the sun. The resulting force density is:

$$\frac{(1.83)2660}{3(10^8)} = 16.2 \mu\text{N/m}^2$$

as the shadow line sweeps across the mirror surface, with a speed equal

to the component of orbital velocity normal to the sun line at that moment:

$$6.52 \sin 37^\circ = 3.98 \text{ km/s}$$

Since the α satellite is only 1 km across, the transition from total darkness to total sunlight takes only 0.25 s, a truly abrupt dawn. Still, the total force of sunlight pressure on α is only:

$$16.2(10^{-6}) \pi (500)^2 = 12.7 \text{ N}$$

easily cancelled dynamically as it comes and goes by a controlled combination of propulsion and active structural deformation. We expect the thermal effects associated with sudden dawn and dusk to dominate the control challenge.

References

- AGARD- Guidance and Control Techniques for Advanced Space Vehicles NATO Advisory Group for Aerospace Research and Development, Conference Proceeding 350 (Specialised Printing Services, 1984).
- M P Ananda, W L Sjogren, R J Phillips, R N Wimberly, B G Bills- [op cit Chapter 6].
- Anthony K Amos- "Air Force Basic Research in Dynamics and Control of Large Space Structures" Proceedings of the 56th Shock and Vibration Symposium Monterey (Oct 1985).
- Satya N Atluri- "Space Structures" Aerospace Engineering Seminar at University of Maryland in College Park (27 Jul 1987).
- J N Aubrun, C Z Gregory, M G Lyons, R L Kisut, A A Woods Jr- Vibration Control of Space Structures VCOSS A: High and Low-Authority Hardware Implementations Final Report AFWAL-TR-83-3074 (Jul 1983).
- H Baruh- "Actuator Failure Detection in the Control of Distributed Systems" Proc 5th Symposium: Dynamics and Control of Large Structures Blacksburg VA (Jun 1985).
- R R Berggren & G E Lenertz- Feasibility of a 30-meter Space Based Laser Transmitter NASA CR-134903 (Oct 1975).
- W E Berkman- "Use of the Laser Interferometer for Position Feedback" Interferometry SPIE Vol 192 (1979).
- Carl Bowin- [op cit Chapter 6].

Burleigh Instruments, Inc- product technical information,
(Fishers NY 14453) (1987).

Jean-Pierre Changeux- Neuronal Man - The Biology of Mind
(Pantheon, 1985).

J C Chen & J L Fanon- "Stiffness Control of Large Space
Structures" NASA 2nd International Symposium on
Aeroelasticity and Structural Dynamics (1985).

Patricia Smith Churchland- Neurophilosophy: Toward a Unified
Science of the Mind-Brain (MIT, 1986).

John B Dahlgren & Lawrence W Taylor Jr- "Spacecraft Control
Research at NASA" in [AGARD 350].

John G Daly- "The Nd:YAG Laser Rangefinder/Designator"
Scientific and Engineering Applications of Commercial
Laser Devices SPIE Vol 610 (1986).

L Davis, N E Buholz, C W Gillard, C C Huang, W M Wells III-
"Structural Alignment Sensor" in
[Naumann & Butterfield, 78].

R P DePaula, E L Moore, R S Jamieson- "Overview of
Fiber-Optical Sensors" Technical Support Package from
NASA Tech Briefs Vol 11 No 7 Item 84 (Aug 1987).

H Dougherty, A M Brady, L F Reschke, K Tompetrini, S P
Weinstein, J Blondin, H Kleine, J Rosenberg- "Noise
Characterization and Minimization of a Precision
Gyroscopic Rate Sensor" in Proceedings of the 9th
IFAC/ESA Symposium P Th L M van Woerkom, ed
(Pergamon, 1983).

- J L Fanson & Jay-Chung Chen- "Structural Control by the Use of Piezoelectric Active Members" in NASA/DOD Control/Structures Interaction Technology 1986 (Robert L Wright, ed) NASA CP-2447 (1987).
- A G Guy- Essentials of Material Science (McGraw-Hill, 1976).
- R T Haftka & H M Adelman- "An Analytical Investigation of Shape Control of Large Space Structures by Applied Temperatures" AIAA Journal Vol 23 No 3 (Mar 1985).
- "Effect of Sensor and Actuator Errors on Static Shape Control for Large Space Structures" AIAA Journal Vol 25 No 1 (Jan 1987).
- William K Hartmann- [op cit Chapter 2].
- Harold M Hastings & Stefan Waner- "Neural Nets on the MPP" Frontiers of Massively Parallel Scientific Computation NASA CP-2478 (Sep 1986).
- Michael Hercher & Geert Wyntjes- "Precision Angle Measurement with a 2-Frequency HeNe Laser" Design of Optical Systems Incorporating Low Power Lasers SPIE Vol 741 (1987).
- Robert J Herzberg- "Controls Technology: the Key to Future Space Systems" in [AGARD 350].
- Paul L Hewitt (ed)- Modern Techniques in Metrology (World Scientific, 1984).
- V O Hoehne- "AFWAL Space Control Technology Program" (Dec 84) in [Boyer, 85] [op cit Chapter 7].

Leonard W Howell Jr- A Stochastic Model for Particle Impingements on Orbiting Spacecraft
NASA Technical Paper 2550 (Jan 1986).

D M Hunten, L Colin, T M Donahue, V I Moroz-
[op cit Chapter 6].

Anthony M Johnson- "Modular Hand-Held Eyesafe Laser Rangefinder" Scientific and Engineering Applications of Commercial Laser Devices SPIE Vol 610 (1986).

Marshall H Kaplan- Modern Spacecraft Dynamics & Control
(John Wiley & Sons, 1976).

Ernest W Kent- The Brains of Men and Machines
(McGraw Hill, 1981).

T Lange & H Holzach- "Integrated Sensor and Actuator System for Flexible Space Structure Experiments" ESA 2nd European Space Mechanisms and Tribology Symposium
(Dec 1985).

Lansing Research Corporation- product literature
(Ithaca NY 14850) (Jul 1984).

Nick LaPrade- "Aerospace Highlights 1987: Communications"
Aerospace America Vol 25 No 12 (Dec 1987).

R A Laskin & S W Sirlin- "Future Payload Isolation and Pointing System Technology" Journal of Guidance, Control and Dynamics Vol 9 No 4 (Jul/Aug 1986).

Daniel Malacara (ed)- Optical Shop Testing (John Wiley, 1978).

Julie Ann Miller- "Chips on the Old Block" Science News
Vol 129 p 408 (28 Jun, 1986).

Mortimer Mishkin- "Our Dual Memory Systems" Engineering
Colloquium, NASA Goddard Space Flight Center
(18 May, 1987).

F M Mottier- "Microprocessor-Based Automatic Heterodyne
Interferometer" Advances in Optical Metrology SPIE
Vol 153 (1978).

NASA Tech Briefs- "Solidifying Bi/MnBi at Low Gravity" p 286
(Spring 1983).

- "Producing Low-Oxygen Samarium/Cobalt Magnet Alloy" p 60 (Jul/Aug 1987) (a).
- "Six-Axis Superconducting Accelerometer" p 72 (Jul/Aug 1987) (b).
- "Thermally Activated Driver" p 62 (Sep 1987).
- "Three-Axis Superconducting Gravity Gradiometer" p 87 (Nov/Dec 1987).
- Neural Network Developments (entire issue) Vol 12 No 1 (Jan 1988).

E C Naumann & A Butterfield (eds)- Large Space Systems Technology Vol 1 NASA CP-2035 (1978).

John Neff- "Optical Computing" delivered at Aerospace: Century XXI AAS 33rd Annual Meeting, Boulder (Oct, 1986).

R S Neiswander- "Inflight Optical Measurement of Antenna Surfaces" in [Naumann & Butterfield, 78].

- "Surface Accuracy Measurement System for Deployable Reflector Antennas" AIAA Paper 79-0937 (1979).

Matti Paunonen- Studies on the Metsaehovi Satellite Laser Ranging System Report of the Finnish Geodetic Institute (Helsinki, 1982).

Vincent Pisacane- "Aerospace Highlights 1987: Computer Systems" Aerospace America Vol 25 No 12 (Dec, 1987).

W J Rowe- "Prospects for Intelligent Aerospace Structures" AIAA Paper 86-1139, AIAA/SOLE 2nd Aerospace Maintenance Conference, San Antonio (May 1986).

J R Sesak, M J Gronet, G M Marinos- Passive Stabilization for Large Space Systems NASA CR-4067 (Apr 1987).

H Sielmann & A Balsarowicz- "Electromechanical Tuning Element With Extended Range for Dither Stabilization of Lasers" Review of Scientific Instrumentation Vol 55 No 10 (Oct 1984).

Stepan S Simonian- "Synthesis of Discrete Passive Vibration Dampers" Structural Dynamic Testing & Analysis Aerospace Technology Conference and Exposition, Long Beach CA SAE Paper 851935 (Oct 1985).

V N Smirnov, O L Vaisberg, D S Intriligator- "An Empirical Model of the Venusian Outer Environment 2: The Shape and

Location of the Bow Shock" Journal of Geophysical Research Vol 85 No A13 (30 Dec 1980).

P Studer, A Baz, R Sharma- "Actuators for Actively Controlled Space Structures" Acquisition, Tracking, and Pointing SPIE Vol 641 (1986).

Stephen Z Szirmay- "Stability and Control of Future Spacecraft Systems" AIAA Paper 79-0864 (1979).

James R Wertz- [op cit Chapter 6].

Stephen Wolfram- "Cellular Automaton Supercomputing" Frontiers of Massively Parallel Scientific Computation NASA CP-2478 (Sep, 1986).

Gordon R Woodcock- [op cit Chapter 7].

W C Yager- "The Precision Self-Metering Structure" in [Naumann & Butterfield, 78].

CHAPTER 9

SPACECRAFT SYSTEMS

Chapter Abstract - Active structural trusses consist of C/Mg composite tubes and nodes, adjusted by PZ strain films and aluminum thermal actuators and monitored by embedded fiberoptic sensors. Reactive metal foams, titanium fittings and conductive tethers serve other specialized structural functions. Nuclear reactors power the fleet with static conversion; the smaller craft use thermionic plants while the largest use thermoelectric. Large diameter annular momentum-control devices perform attitude trim maneuvers. Rim desaturation is propulsive and by tether management, which also supplies large and constant control torques. Xenon ion engines, variously sized and ganged, comprise the propulsion plants of the spacecraft. People are forbidden to approach the finely tuned ships' delicate hardware; a subfleet of robot manipulator craft, under interactive control of the fleet intelligence, effects all inspections, preventive maintenance and necessary repairs.

In this chapter we examine systematically the feasible options for providing structural, power, thermal control, attitude control, propulsion, and maintenance services to the satellites comprising the planetary laser fleet. Each section, after reviewing available technologies, then chooses the method or combination of methods best suited to the fleet's needs in that area, highlighting departures from standard application where appropriate.

Structure

The basic role of major bus structure varies throughout the spacecraft fleet. The small satellites like the Transducers and Station 16 require bus structure primarily as a way of connecting their power plants and payloads to their AMCDs. Intermediately sized craft like the Rings and Station 18 essentially are AMCDs carrying a few attached objects, so their main structure is the AMCD armature. Finally, the really big craft comprising the planetary resonator itself require major structure to serve both of these functions secondarily as well as the primary job of providing a precisely reacting "ground" for mirror actuation. In all cases, we have seen the vital need for these structures to be intelligently adaptive, given the uncompromising optical requirements and enormous dimensions of our laser communication system.

Any contemporary, extensive structural optimization (analytical or more typically numerical) of the fleet craft would be an inappropriately detailed exercise, considering the uncertainties propagating down through every level of this

feasibility study. Even the actual reference systems-outlines of Chapter 4 are so approximate as to be fairly insensitive to the finesse of a rigorous structural design. By moving the optical accuracy dilemma under the aegis of active control as we did in Chapter 8, we stripped just about all the intrinsic glamour from the fleet's raw structural performance. Methods for load transmission, in fact, represent one of the most straightforwardly solvable problems posed by these spacecraft. What we can and will do here is select the major structure systems arguably best fitted to the tasks listed above, and select as well some material systems which can manifest them.

Even for most terrestrial applications, and certainly for optical applications, the strength of a structure (its ability to carry load without coming apart) is far less limiting than its stiffness (its ability to carry load without excessive distortion). So far in the history of structural economy, geometry has been a material's only ally in resisting deflection. Thus typical structural members are not only made of high-elastic-modulus materials such as metals (although other properties of metals like workability and fracture toughness are critical advantages also), they exhibit shapes maximizing their stiffness per unit mass. Wide flange sections, deep or box beams are favored for bending, closed hollow sections for torsion, and "short" tubular or laterally supported bars for compressive stability. When excellent materials are also configured optimally, uniquely high strength/mass performance results, as in the case of supersonic jet aircraft.

The mathematical models of structural mechanics reveal quite well where a structure's material is needed most for a given load pattern, so that clever geometry precludes using excessive material. Three-dimensional trusses demonstrate most clearly the extreme lightweighting achievable. Such systems

reduce bending and twisting to simpler compression and tension by constraining loads to follow axial paths through thin members. Failure modes are then limited practically to local member performance. Again, by combining such geometrical economy with superior materials, rivers can be spanned and large space stations can be erected from small launchable packages.

The active stiffness control featured by all the fleet structures represents the next major step in structure system evolution (Figure 9-1). Closed-loop strain actuation thus joins geometry and material development as a new enabling tool for superior performance. Truss members under active control, for example, can use yet less material even than the improbably slender bars we customarily see planned for orbital frameworks. Plates with segmented, active surface strain films can change macroscopic shape with microscopic resolution, to achieve desired transient geometries or to cancel environmental effects. Locally applied forces and torques can be isolated and neutralized before they propagate throughout an extended structure. Most importantly, a flimsy structure only one "bay" thick can be made to act as rigid as one which is hundreds of times thicker and thousands of times more massive, even if it is kilometers in extent. A few members working together can appear as infinite a "ground" as necessary to react actuator forces with nm precision.

Such performance does not come for free, of course. Just as the price of clever geometry is complicated analysis and difficult manufacture, the price of active structure is power consumption and processing ability. As we will see in the next section, our large satellites are already constrained by their mirror control needs to use nuclear power, so our system cost for powered structure is "marginal" only -- ACOSS changes the size but not the type of power plant we specify. As we

know from Chapter 8, however, information-processing complexity does not scale up so simply. Clearly the immense computational burden of our system derives primarily from its monitored and controlled structures. What can be said is that the fleet's ability to adapt to both transient and lingering conditions is not merely a performance advantage, it is a prerequisite. By augmenting good materials with both optimal geometry and a sensory/motor intelligence (as natural selection has been doing with evolving life for hundreds of millions of years), we enable new performance standards.

Trusswork is the obvious system choice for our large spans. Considering inevitable component failures, multiply redundant load paths (making it more than just one bay deep...) are strictly necessary. Although the optimal mix of active and passive members would depend on detailed study, we posit standard truss pieces as self-contained active units. A tubular section both optimizes structural efficiency and provides a service chase for the actuator equipment, and system power and intelligence lines. Since the tube wall is thin compared to its diameter, PZ strain films applied to the inside are almost as effective as external films would be, and are modestly protected as well from space particle fluxes by the wall thickness. A segmented array of strain films can effect subtle member extension, contraction, bending in any direction, and vibration damping. Careful actuator distribution along the member length, based on detailed predictions of strain energy concentrations given these various distortions, can avoid "wasteful" control authority [Simonian, 85].

Processing for the first level verification loop of displacements, and for temperature calibration, is contained within the unit as well. Fiberoptic temperature sensors (decoupled from the member structural strain) are distributed along its length, both outside it and within the chase.

Additional fiberoptic strain sensors are embedded in the wall material itself. Coaxial, modular service connectors at the member ends (being an interface, the most probable source of member failure) serve the triple function of transmitting loads to the joints, supplying the member with its power, command and telemetry lines, and channeling regional power and intelligence through to other nodes.

The predominantly axial loads in these members lead as naturally to fibrous composite fabrication as does the desire for low mass, a conclusion borne out by current space station plans [NASA TB, 8705]. While a variety of material choices exists for both fibers and matrix, we select as the reference tube material a carbon fiber / magnesium matrix (C/Mg) composite with fiber volume fraction 0.49. This material has been developed particularly for space applications requiring high specific stiffness and desirable thermal properties [McDanel's et al, 86]. First of all, carbon fibers have a negative longitudinal coefficient of thermal expansion (CTE) [NASA TB, 80f-a], so embedding them in a matrix with positive CTE can yield a thermally stable layup of near-zero composite longitudinal CTE [Remondiere et al, 85]. But epoxy matrices, having low thermal conductivity, allow destabilizing transverse thermal gradients to remain in differentially heated members (as for instance those exposed to sunlight). A high thermal distortion coefficient (transverse thermal conductivity κ_T divided by longitudinal CTE α_L) implies that dimensional changes are relatively small, and occur and equilibrate quickly; this criterion favors metal matrix composites (MMCs) by a wide margin (a minimum factor of 33) over those with glass or epoxy matrices [Remondiere et al, 85].

C/Mg surpasses C/Al, the other "light metal" MMC, in both thermal stability and specific stiffness, and at lower fiber volume fraction. Making the material is not trivial, but

several fiber-yarn wetting techniques are being refined, including magnesium plasma vapor deposition (PVD) of the carbon, and an alkali metal intercalation pretreatment, both of which would be simplified by vacuum space processing. Because magnesium oxidizes rapidly [Kalpakjian, 85], we would expect some gradual changes in surface properties of those members used on the resonator satellites around Venus. However, the members will be externally finished with a low-absorptivity, high-emissivity coating to moderate their environmental heat exchange (as discussed later in this chapter), and we expect this to protect the metal sufficiently from long-term orbital oxygen flux. In any case, magnesium is less susceptible than epoxies to space and radiation degradation.

Finally, we note other advantages of MMCs [McDanel et al, 86] for our application. Metallic (conducting) members can avoid the disruptive problem of differential static charging over time, as the spacecraft move through the solar wind and any regions of magnetically trapped particles. This is particularly crucial considering the damage suffered by PZ materials from arcing. Also, the intrinsic fracture toughness of metals confers greater impact resistance on MMC than competing composites demonstrate. For us this means that a "large" (1 mm, which can easily carry 2 kJ of kinetic energy) meteoroid has a higher chance of penetrating rather than shattering completely a truss member that it hits. The control intelligence can compensate better for crippled members than for obliterated ones. Finally, techniques have already been patented [Aerosp Am, 87] for MMC capacitative discharge welding. This should decrease the cost and increase the reliability of, for instance, attaching temperature sensor mandrels to MMC members. It also suggests the possibility of distributing simply hard-welded, passive truss nodes throughout the structure. (Just as with the truss members themselves, a

percentage of passive joints would reduce both cost and complexity if control accuracy could be maintained.)

The reference truss node fitting, though, is also a sophisticated active unit which does much more than join the active members mechanically. Contained within the node core is second-level processing which coordinates commands to, and telemetry from, the members radiating from it, communicating with its neighbor nodes through the next, third-level. Bonded between the core and each member socket is a series thermal actuator, which actively sets the long-term member length bias about which occur the member's higher frequency PZ changes. These thermal expanders are segmented aluminum sleeves, electroresistively heated and radiatively cooled, instrumented fiberoptically for continuous calibration.

The nodes also contribute some passive vibration damping, which as noted in Chapter 8 can slightly reduce the structure's active control burden. Major passive attenuation is not achievable, however. The only viscoelastic (VE) materials known which can act in structural series with members are cross-linked polymers [Soovere & Drake, 85], which would of course be prey to space degradation. More critically, these materials dissipate negligible energy unless strained more than our active system can allow. That is, optical performance requires the high-frequency PZ actuators to cancel immediately any motions which would be large enough to make VE constrained layers useful.

However, magnesium exhibits damping comparable to alloys of copper and zinc ($\zeta \approx .03$), much greater than conventional structural metals. Crystal defect dislocation is the dissipative mechanism [Soovere & Drake, 85], so naturally stiff carbon fibers limit composite damping substantially. Still, with $\zeta = .001$, C/Mg provides the greatest damping of any

aerospace composite, even at a much higher fiber volume fraction than we are using. Thus our truss members provide a marginal damping bonus. The nodes provide even more, because their structural shells are of discontinuous C/Mg. Reinforced by chopped fibers, they display most of the enhanced strength but not all the enhanced stiffness of continuously-reinforced layups, allowing better damping [Misra, 86]. The chopped C/Mg material is optimal for fashioning fittings (like the truss nodes) because it is easily cast or pressed using powder metallurgy processing. Fine machining must be done chemically because of the abrasive carbon inclusions. Predictably, our nodes will be fashioned with embedded sensors to monitor their strain and temperature.

The reference truss system we have just defined can, in one guise or another, satisfy the structural needs of the fleet's large spans as enumerated at the beginning of this section. The combination of material, configuration and control we have worked out can be scaled throughout a range of load and dimensional requirements, to be used in the worst chemical, particle and radiation flux environments our craft will see. The system can alter shape by meters (over long spans) at resolutions of nm, with μm strokes at hundreds of Hz. System accuracy is design-limited by the amount and quality of additional, external state sensing specified, and the speed and sophistication of distributed control processing built in. Operationally, accuracy is limited by the frequency and extent of clustered local failures.

The node material, its fabrication process and weldability are directly applicable to the myriad connection fittings, apart from trusswork, which the fleet needs. Three other specific types of structural elements deserve close attention, however.

We can presume that for structure systems smaller than a certain dimension (knowable only after detailed tradeoff study), the discrete-member active truss approach to optically accurate performance would be less efficient than some continuous system, such as an active plate or active beam. We acknowledge this possibility in the reference designs for our smallest satellites, the Transducers and 16, where we specify monolithic frames shape-tuned by segmented PZ surface films. Such structures should be light and stiff, although volumetric bulk is not critical (their smaller relative total size presents a reduced space particle flux cross section in any case). Particularly for unmanned, space-based environments without reactivity dangers, materials of interest for this use are ultralight reactive metal foams (RMF) [Cocks, 84].

These are alloyed of metals whose high oxidation potential precludes atmospheric use, but whose low densities make them attractive for structural purposes in vacuum. When foamed in the molten state, their density decreases even further (a material improvement) as their volume increases (a geometrical improvement). Enhancement of structural stiffness results, since even if the material stiffness decreases by the volume ratio, the sectional area moment of inertia increases in general by the square of the area ratio, which is greater. Large, homogeneous foamed metal castings would be achievable in microgravity. A well-studied RMF alloy is Mg(85 %)-Li(14 %)-Al(1 %), which expands an order of magnitude volumetrically when heated together with the admixed foaming agent BaH_2 , yielding a density of only 135 kg/m^3 .

More work is of course required to characterize such materials in detail, and study sensor-fiber and actuator bonding techniques. But it is known that the Mg-14Li-1Al alloy "possesses exceptional resistance to hypervelocity impacts by small particles" [Cocks, 84]. Conventional micrometeoroid

bumpers are spaced multilayer constructions; the first layer vaporizes an incoming particle, and the subsequent layer stops the resulting fines. We could therefore reason that this alloy when foamed would act as a multiply-spaced layered barrier, for yet greater penetration resistance. Also, any alloy of these three metals certainly possesses good thermal conductivity; we would expect this to decrease in proportion to the foamed area ratio, but the specific thermal conductivity must remain extremely high. We will assume RMF fabrication for most of the fleet's non-truss structures, protected from the weak oxygen flux seen by the orbital craft by a plasma-sprayed coating.

Next, there are a few structural applications, especially around the nuclear reactors, their conversion equipment and radiators, for which high strength and stiffness and (of course) low mass are required, but with the additional stipulations of radiation resistance and low thermal conductivity. Because low κT composites are radiation-susceptible, the natural material choice is titanium. Among the engineering alloys, titanium has the lowest thermal conductivity [Kalpakjian, 85]. Its well-known specific strength and stiffness can thus be used for support structures that should also minimize conductive thermal contamination.

Certain titanium alloys (such as Ti-6Al-4V) have the further advantage that complicated geometrically stiffened sheet shapes can be easily formed superplastically [Kalpakjian, 85]. That is, at the proper temperature and extremely low strain rates (down to $\sim 10^{-4}/s$), the alloy will deform like molten glass with elongations as high as 2000 %. When properly tooled, this process yields high-precision finished pieces in a single operation. Used currently for stiffened airframe panels, we expect it to simplify production of many structurally optimized fittings and elements for the fleet.

Finally, as justified in Chapter 8 and implemented in the reference design of Chapter 4, our fleet requires space tethers. To date only extremely short tethers have been used in space, but a 100 km tether experiment planned for STS deployment will soon open in earnest this area of operations, surely destined for great achievement. The field of space tether theory and applications is already developed enough, and our uses for tethers mundane enough compared to serious proposals, to warrant including such devices in our fleet.

The gravity gradient stabilizes "vertical" tethers. The mass below the CG's orbit tries to travel faster, while the mass above it tries to travel slower; the resulting tether tension (equal to the gravity gradient force) constrains them instead to move together. Because of the inverse-square attenuation of gravity, a tethered system's CM and its CG are not the same, but the difference is noticeable only for extremely long tethers. For "short" tethers (< 200 km) the tether mass itself is negligible compared to the masses it can hold, so the structural efficiency of a constant cross section is indistinguishable from that of the tapered cross sections necessary for long systems [Baracat & Butner, 86]. Our tethers are just a few kilometers long, enough to assure their stability generally [Lemke et al, 87].

No work has yet been done on the behavior of tethered systems in the presence of large local masses, the situation at Station 1. To first order, however, we note that the gravitational attraction by the large satellites of the smaller tethered ones (which would tend to slacken their tether) is of order 10 N (Appendix A8-3), an order of magnitude smaller than the tether tension (Appendix A4-4). We will therefore assume stability for this study. Three types of vibration predominate for tethers in space: in-orbit-plane libration (swinging) with a period of $\sqrt{3}\Omega$ (where $\Omega \equiv$ the orbital

rate), out-of-plane libration with a period of 2Ω , and transverse "vibrating string" behavior [Arnold, 87]. A tether will go slack if its in-plane libration exceeds 65° or its out-of-plane libration exceeds 60° (both of which are far in excess of what our inter-aiming tolerance can permit); libration can be controlled by timed, active "reeling" [Baracat & Butner, 86]. In general, this is best accomplished by having the satellites crawl along a permanently deployed tether, rather than having them reel it in and out [Glickman & Rybak, 87]; $l\delta$ is our active crawler. Provision must be made for tuning the location of the tether attach point with respect to a spacecraft CM, to avoid (or to use) attitude torques which can be several orders of magnitude larger than any others available; all the tether anchors in our fleet feature 2 DOF actuation for this purpose.

Our reference design includes two types of tethers: an altitude and attitude compensation tether for $l\gamma$, using the fleet warehouse and repair shop as counter-mass, and a power tether linking $l\delta$ with $l\epsilon$. The first of these poses mainly a structural design problem because of the enormous mass of $l\gamma$. Since we have not designed that satellite in detail, we will not dwell on its tether, except to note that the quest for high-strength tether materials has ranged so far from aluminum, through steel and titanium, to Kevlar, and expects to include advanced polymers, then graphite and silicon carbide [Baracat & Butner, 86]. Some form of tether system to sustain 35 kN (about what this job would call for) is eminently feasible -- a Kevlar cross-section less than 2 cm^2 could do it with the standard tether safety factor of 3.5.

The tether between $l\delta$ and $l\epsilon$ is much more interesting, because it must transmit about 335 kW (Appendix A4-3) as well as sustain 130 N (Appendix A4-4). The electrical constraints of a power tether usually dominate its structural ones, with

the minimum sectional area being limited typically by its need to radiate resistive heat to space [Martinez-Sanchez & Hastings, 87]. The baseline material for such tethers is aluminum, but particle impact safety considerations favor developing tethers with non-structural conductors wrapped helically around a core of high elastic strength [Scala et al, 87]. A promising new composite tether design consists of Kevlar microstrands impregnated with copper for conductivity and overcoated with nickel for chemical protection in low orbits [Orban, 87]. Although we base our system mass on the primitive aluminum type, we expect that our tethers would indeed be the more advanced type. In fact, the redundant optical telemetry and intelligence fibers the tethers must carry can be incorporated easily into braided constructions [Bevilacqua et al, 87].

Unlike power tethers in the literature, ours does not close its circuit through plasma contactors to space; we need separate "up" and "down" conductors, which might best be served by separate tethers. This arrangement, which we specify, has two additional advantages. First, structural redundancy (practically precluding any risk from catastrophic meteoroid severing) is automatic with physically separated tethers. Second, such spacing allows inter-tether cross bracing with light trusses, which can limit the tethers' lateral vibrations [Baracat & Butner, 86].

Power

Based on their mirror-actuation power needs alone, our large fleet satellites are limited to a power system menu of one: nuclear. Considering the power needs of their

intelligently active structures, the intermediate-sized craft also require nuclear power plants. And satisfying the propulsion needs of the smallest members of the fleet (the Transducers) is simplified too by specifying nuclear power.

While no spacecraft engineer would question for a moment the choice of nuclear power for our advanced, automated, power-hungry fleet, we now justify briefly such a categorical decision. Providing MWe with chemical systems is certainly feasible [Mil Space, 84], but for such short times (minutes) due to reactant consumption and at such high resupply cost that only burst weapons could benefit. The alternative, non-expendable source is solar.

Solar radiation is almost twice as strong (2.6 kW/m^2) at Venus as at Earth, but even the most advanced types of solar photovoltaic converters achieve less than 20 % efficiency [NASA TB, 80s & 8701]. So each MWe produced would require about 2000 m^2 of collector area under optimal conditions. The specific power achievable with advanced solar systems is 15 to 25 We/kg [El-Genk et al, 85]. Not all of that power would be available for the payload, though. First, the large expanses of collector would themselves require active dynamic control just like the rest of the spacecraft, increasing system complexity and consuming more power. Also, the 29 % of each orbit eclipsed by Venus (Appendix A8-7) would necessitate a heavy, complex power storage system, with attendant propulsion- and attitude-derived power consumption, and at least a 50 % increase in collection capacity to charge it. Finally, system losses and collector degradation (from connection failures and radiation damage) subtract even more from the available power. We might roughly estimate a 5000 m^2 collection area for each MWe of online payload consumption. Solar dynamic systems, while capable of much higher overall efficiencies, suffer from both the same power storage requirement and their vibration

loading of the spacecraft, a disturbance we would rather avoid.

Compared to the alternatives, nuclear power systems are extremely compact, necessarily radiation-hardened, easily unit-replaceable when depleted, and provide steady power reliably for years at a time. Especially for load levels above 100 kWe, the power density of nuclear sources makes them preferable for space applications, and we will power the entire laser fleet with a family of nuclear reactors. Partly because a few types of nuclear power plants have already flown in space, and particularly because of the contemporary SP-100 program to develop a small flight-qualified reactor in anticipation of lunar, propulsion and defense missions, substantial literature exists concerning space nuclear power systems. Since we review here only enough salient information to aid system selection, curious readers should consult the referenced work for further details.

The two kinds of nuclear plants are reactors and radioisotope thermoelectric generators (RTGs). RTGs are used on every deep space probe flown or planned, as well as many defense satellites, because they produce continuous power for decades, are extremely simple with no moving parts, and satisfy a small craft's typically modest power needs. They use the alpha decay of ^{238}Pu to generate electricity with direct static thermoelectric conversion (which we will review later); waste heat is typically rejected to space through simple finned radiators. Being modular, RTG units are easily ganged for greater system power, but because they require a large amount of heavy isotope, are generally limited to a total output less than 10 kWe [El-Genk et al, 85]. A next-generation unit will achieve almost 10 W/kg system power density, packaged in 500 W units [Chipman, 85]. RTGs would prove useful for powering some of the fleet's non-contacting systems (like the secondary Ring mirrors) should they ultimately require

dedicated sources. But since RTGs cannot practically power even the smallest of our entire craft, we will not focus on them further as major systems.

Nuclear reactors for space power systems typically generate electricity with heat from fissioning highly enriched ^{235}U fuel [Bunch, 85], achieving power densities between 40 and 55 We/kg [El-Genk et al, 85]. The four classic subsystems are: the reactor core, its shields, a power conversion plant, and a radiator to reject waste heat to space. Due to the severely mass- and size-crippling 20th century launch limitations, space reactor development has departed markedly from terrestrial reactor design in an attempt to maximize reactor efficiency and hence compactness. Thus present schemes presume a fast-spectrum neutron cycle, in which fission-product neutrons largely contribute directly to the chain reaction without first being slowed (to "thermal" neutrons) through scattering. The chief advantages are a smaller core and a higher operating temperature (which improves conversion and thermal rejection efficiencies), but those same higher temperatures introduce a host of practical problems, some of which have yet to be really solved even despite application of a startling zoo of exotic materials [Buden & Lee, 85].

We must beware this current trend toward small, light and complicated space reactors, for two reasons. First, a planetary laser system could not even be built in a climate of quite so parochial limitations as absolute size and mass ceilings (Chapter 11). So the extreme assumptions underlying all current space reactor work are for us essentially moot, being artifacts of a uniquely constrained moment in history rather than of intrinsic physical laws. Our power plants can afford to be a little bigger, heavier, less efficient and therefore much less amazing than those we will review

momentarily. So even without invoking unpredictable future breakthroughs, our reference designs, by using conservative parameters, will include generous specification margins. For the price of slight clumsiness compared to what is (thought to be) possible, we will purchase vastly more reliability and longevity (especially for thermionic systems), the two features of overwhelming importance for our fleet.

Our second major change from conventional space nuclear planning involves reactor shielding. Great effort has gone into developing absolutely the smallest package to shield (barely and only) the power system and payload from unacceptable radiation fluence, guided of course by the grail of minimum mass. The peculiar "badminton shuttlecock" geometry of the SP-100 (Figure 9-2), which spews hard radiation in almost all directions while operating [Wetch et al, 85], derives directly from such optimization. Figure 9-3 reveals, however, that shield mass dominates overall power plant mass only for small reactors of the SP-100 class. MWe plant mass is dominated instead by the converter equipment, and multi-MWe (if thermodynamic) plant mass completely follows its radiator mass. Given the context, then, not just of the other power plant elements but of our huge vertex stations, the extra material and mass penalties introduced by generous reactor shielding are trivial.

Consequently, in contrast with space reactor design found in the contemporary literature, we will not consider reactor proximity, or shield thickness, extent and orientation to be configuration-limiting parameters. Rather we will categorically accommodate a true, rather than perfunctory, shield for each of our reactors within the plant's mass and envelope budget. In fact, we will specify an asymmetrical 4π radiation shield for all reactors in the fleet. Adopting a policy which we can project to be the eventual standard for

complex space systems, we reason that no power plant should pose an excessively life-compromising hazard to expensive robots (or less likely, humans) performing on-line spacecraft maintenance. Such a shield will naturally be heavier in directions facing the spacecraft itself, since there the absorbed radiation fluence (product of flux and time) is greatest over the mission duration. We extend this consistent design policy even to the smaller craft for which the relative system penalties of increased shielding are not trivial.

A typical reactor core loosely based on the SP-100 shown in Figure 9-4 [El-Genk et al, 85] is a cylindrical close-packed array of finned molybdenum/rhenium (Mo-13Re) heat pipes, each charged with liquid lithium. Occupying the volume between the fins surrounding each pipe are wafers of enriched UN fuel. UN is chosen rather than UO_2 [Cox et al, 85] to avoid lithium reduction of the fuel to metallic uranium (which would melt) in the event of heat pipe rupture. UN is also a more efficient space fuel, being atomically denser. The fuel can be expected to swell several volume percent over the core lifetime (more than UO_2 would), as the small fraction of fissioning uranium is "burned up" into other elements. Since some fission products are gases which must be vented to space, the core should not be hermetic, although UN evolves less gas than UO_2 . A variety of fission flux sensors are being developed for monitoring the criticality of high-temperature space reactor cores [Anderson & Oakes, 85], and quartz fiberoptic temperature sensors have been demonstrated accurate to 2200 K. In particular, the material degradation of fiber glasses from radiation tends to be annealed out continually at high temperatures [Partin, 87].

Enclosing the compact core itself is a set of control drums, each made of the neutron scatterer (reflector) beryllium, but one-third sectorized with B_4C "poison" (whose

boron is 90 % enriched with ^{10}B for neutron absorption). Rotating these drums controls the core criticality by either absorbing fission neutrons, or reflecting them back into the fuel to sustain or increase the chain reaction. In the basic configuration, one end of the cylindrical core is capped by a Be reflector; the other end, where the high-temperature heat pipes exit, is filled by small neutron-reflecting BeO spheres. ZrO_2 ceramic in various forms is used for thermal insulation, particularly to protect the neutron shield from primary core temperatures.

The fission of ^{235}U fuel produces primarily a menu of over 60 isotopic products, as well as neutrons and γ -rays [Peters, 87]. Both α and β particles are produced also, but the former have little penetrating ability, and the Brehmsstrahlung yielded when the latter collide with matter is absorbed by the γ shield. Of the neutrons evolved, slightly more than half will be absorbed by the fuel to sustain the chain reaction, about a sixth will be absorbed by the refractory core metals, about 1 % will be absorbed by the coolant and insulators, and the rest (about a third) will "leak" out of the core [Homeyer et al, 85], to be absorbed by the poison moderators and the neutron shield. LiH (enriched to 92.6 atomic % ^7Li) in a stainless steel matrix makes a good neutron shield; it has a high hydrogen content, low density, and can capture neutrons without evolving more γ -rays [Barattino et al, 85]. However, its low thermal conductivity and melting temperature require design attention to heat dissipation. A layer of high-Z material (typically tungsten) absorbs core γ -rays. By preceding the neutron shield, this dense layer prevents extra γ -heating of the LiH.

The exit path of the core-cooling heat pipes through the shields must be labyrinthine, to prevent radiation escape conduits. Liquid alkali metal coolants are used for

fast-spectrum reactors because of their superior heat transport performance, their low vapor pressure at high temperatures (improving containment reliability over the reactor's life) and their small reactivity cross-sections for fast neutrons [Bailey et al, 85]. Heat pipes are a better choice than pumped-fluid alternatives for several reasons [Koenig, 85]. Their capillary flow principle, reviewed in the next section, avoids flow turbulence vibration, and they require no power source. Self-starting even after long periods of disuse, they transfer heat almost isothermally (with a temperature drop of only 25 K over several meters). Shorter lithium pipes have demonstrated 22.6 kW transfer rates at fluxes of 19 kW/cm² (axial) and 3.6 MW/m² (radial), operating at 1500 K for years. Because the coolant inventory is comparatively small in heat pipes, they suffer less from neutron activation than pumped-fluid systems, and the small amount of helium which is thus evolved does not affect the pipes' efficiency. Finally, the parallel design of a heat pipe cooling system is inherently highly redundant, making core coolant reliability arbitrarily close to unity. Heat pipes interface easily with all of the power conversion alternatives.

Systems to convert the high-grade heat into electrical energy divide into two classes: dynamic and static. Dynamic converters use either single-phase (Brayton or Stirling cycle) or two-phase (Rankine cycle) heat engines to power electrical generators [Westphal & Kruehle, 85], as in terrestrial nuclear plants (except that the primary loop fluid is typically liquid metal, as noted above, rather than water). A largely undeveloped alternative method uses magnetohydrodynamics (MHD) to generate current directly by passing a high-velocity plasma through a magnetic field. Although dynamic systems achieve the highest overall efficiencies, and in fact have demonstrated the greatest longevity so far, they require heat exchanger stages, pumps, ducts, seals and complex moving parts. The high

rotational speeds and fluid flows in such systems constitute a major vibration source. Primarily because we seek to minimize unnecessary vibration forcing of our optical satellites, our reference design will not consider dynamic conversion systems.

Individual purely static plants are eventually power-limited by the practical distance over which heat pipes can transport core heat; however, their radiator mass grows only in proportion to the converter mass as system output increases to that geometrical "packing" limit [Mahefkey, 85]. There are four available methods of static conversion, all of which are scalably modular: alkali metal thermoelectric conversion (AMTEC), thermionic (THI), thermoelectric (TE), and thermophotovoltaic (TPV) [Ewell & Mondt, 85].

AMTEC is at once the most efficient and adolescent of these techniques. Heat drives liquid sodium to the vapor phase across a permselective β'' -alumina barrier (electronic insulator but ionic conductor). Then giving up its heat to the radiator, the vapor condenses to be recirculated. The ionic gradient potential difference powers an external load circuit (Figure 9-5). Because sodium is a conductor, an EM pump with no moving parts can recirculate the liquid phase, but of course the fluid flow itself means that this system is not strictly static, and would in fact be a source of some vibration. No AMTEC system has operated for more than 1000 hr so far. Despite its demonstrated conversion efficiency of 20 - 40 % [NASA TB, 84] then, we will not consider AMTEC technology for our reference design.

THI converters use heat to boil electrons off an emitter surface, which when collected by an anode set up the potential difference to power an external load. To neutralize the space charge effect whereby the electron cloud would inhibit further electron emission, an easily ionized vapor -- typically cesium,

which also reduces emitter burnup -- must bathe the gap (Figure 9-6). THI units have operated at 17 % electrode efficiency for over 5 yr at emitter temperatures of 1970 K [Ewell & Mondt, 85]. Here again, though, we encounter the zealous extremism of current development. To compact space reactor design as much as possible, THI converters are normally anticipated to be in-core, simultaneously taking advantage of the higher temperatures there and avoiding a need for primary high-temperature heat transport. Unfortunately the in-core environment and fuel swelling are tough on the converters, and none has survived more than a few years. As noted above, though, the technology is quite reliable, and eminently suited to our needs, especially in its somewhat less efficient out-of-core format. The high thermal rejection temperatures of THI systems, because of the Stefan-Boltzmann fourth-power radiation law, confer on them a reduced radiator-area advantage over the other conversion methods.

Not surprisingly, the high-performance in-core THI converter scheme results in a different core design [Homeyer et al, 85] from that outlined earlier, one which is not purely static. As shown in Figure 9-7, a typical niobium/zirconium (Nb-1Zr) alloy core vessel contains many close-packed thermionic fuel elements (TFE), which are cylindrical stacks of self-contained electric generator cells. Each cell features a UO_2 fuel slug inside an annular tungsten THI emitter and precisely spaced (0.51 mm) niobium collector, all wrapped by a Nb-1Zr sheath. The core is cooled by pumped alkali (NaK) coolant, which flows at 1 m/s through the tricuspid interstices between TFEs. Molybdenum electrical leads, alumina (Al_2O_3) insulators and Y_2O_3 seals complete the exotic material cornucopia, and control drums surround the core as described earlier. Presuming that the many TFE technology problems currently being attacked [Holland et al, 85] can be solved, such a core could be scaled up to almost 6 MWe and still have

a diameter less than 1 m [Homeyer et al, 85], demonstrating both the unsurpassed power density of nuclear sources, and their relative configuration insensitivity to power level.

TE converters are simply solid-state thermocouples, in which a potential difference is induced by charge migration across an intermaterial junction whose two sides are maintained at different temperatures (Figure 9-8). RTGs use the TE principle, and have demonstrated its reliability in deep space, albeit at only about 7 % efficiency, for many years. TE efficiency is maximized when material junctions have high electrical conductivity σ , low thermal conductivity κ , and a large difference in Seebeck coefficient α . The figure of merit $Z = \alpha^2 \sigma / \kappa$ is traditionally highest for semiconductors, but work on two new material classes, the rare-earth chalcogenides and the boron-rich borides, looks promising [Wood, 87]. TE powerplants would in general represent the most conservative, least sophisticated option for our fleet.

The TPV principle heats incandescent elements to a temperature whose peak emission is matched to the electron band-gap of surrounding photocells, which then produce electricity (Figure 9-9). It is conceptually the simplest, albeit right now the most intractable, of the static systems. Although capable of short-life conversion efficiencies up to 30 %, it suffers from the worst materials development problems of all, requiring radiation-hard high-temperature photocell materials, Level IV (>1500 K) heat pipe technology, and such low-temperature thermal rejection that its radiator area/We needs to be about five times greater than its nearest competitor conversion method, TE [Ewell & Mondt, 85]. Keeping in mind that unpredictable future breakthroughs might make TPV the technology of choice for efficient, truly static conversion, we will nonetheless not consider it further.

We will use both THI and TE systems. Because of the relatively small radiators they require, thermionic converters will be useful for the smaller satellites whose formations redirect the laser beam extensively, namely the $l\delta - l\epsilon$ pair and the L1 and L2 stations. For the reliability and longevity reasons cited earlier, most of these converters will be the out-of-core type. However, we specify in-core THI converters for the Transducers. While not strictly essential, power plant compactness is desirable for such small satellites. Also, their mechanical modulators represent single-point failure risks already, so the maintenance and replacement we might expect anyway for the Transducers mitigate the system liability of a shorter-lived power plant. In fact, the small relative cost and changing mission requirements of these comparatively tiny but critical craft both point to the wisdom of maintaining a set of spares on standby.

The large resonator satellites, though, will be maintained on-line in most cases, so the extreme reliability, longevity and modular maintenance simplicity of advanced TE converters make them most appropriate for those rather large reactor plants. Chapter 4 shows that even the large radiator area required easily fits into these huge craft, with the total power plant mass comprising about 20 % of the spacecraft mass.

The reactors provide a steady source of heat, with which their converter systems provide a steady source of electrical power. During nominal optical operation, we may presume that fine control consumes generally a steady amount of power, averaged both over time and the spacecraft dimensions. But there will be times when the consumption pattern changes fairly drastically, such as when the Transducer and Ring interrupt fine control to reorient or reposition themselves, or when the active structure of the large resonator satellites is under

construction, or must damp out quickly the effects of any large transient loads caused by catastrophic debris impact. No power plant is complete, therefore, without power conditioning equipment to even out the changing load. The best method of regulation depends upon the exact converters used in the plant, but for the TE systems of our large satellites, dissipative shunt regulation (a controlled variable load across the "user" load) is best [Kirpich & Yadavalli, 87]. It requires the smallest marginal radiator area, can be located safely away from the nuclear core, follows the user load well, and does not disturb the reactor's steady operation.

Thermal Management

Rejecting waste heat from the power converter assemblies presents the major thermal control challenge of the fleet, although other thermal management tasks can be grouped as follows: removing low-grade waste heat from power regulation and distribution equipment and datonics, stabilizing thermal behavior of the structure systems, and cooling cryogenic sensors. Spacecraft waste heat must ultimately be radiated away, of course, since in space there is nothing to convect in, or conduct to.

As explained earlier, compactness and lightness are not the drivers of our power systems. Our selection of static converters, based on a desire both to avoid bus vibration sources and to facilitate modular maintenance over long mission times, has a direct effect on our choice of radiator systems. Because the dynamic converters we are not using are extremely compact heat sources, invariably cooled by pumped fluid, their heat is most naturally rejected using compact, dynamic

radiators. Thus, such fascinating sources of vibration as belt radiators [Feig, 85], viscous filament radiators [ibidem], liquid droplet radiators (LDR) [Mattick & Hertzberg, 81], rotary disk radiators [Elliott, 85], and two-phase rotating bubble membrane radiators (RBMR) [Webb & Antoniak, 86], while common in discussions of space nuclear power, will not appear in our reference design for the Venusian laser. We will encounter them for another application entirely, however, in Chapter 10.

The very modularity of the static converters makes them distributed systems, so that we are led to radiation methods which are as distributed. Having spread the high-grade heat around the converter assembly, after all, we should reject the low-grade heat then and there, rather than re-collect it with another thermal utility. TE converters typically use the cold shoe of the thermocouple itself as the radiator surface. No additional hardware (or complexity, or mass...) is then required, but the desire for a unity space view factor from the radiating surfaces constrains the overall converter geometry to simple curvature, as demonstrated by the convex SP-100 configuration. THI converters also can use their anode directly as the radiator, particularly efficient since their rejection temperature is higher than thermoelectrics', although this constrains their layout as noted for the TE converters.

To satisfy specific spacecraft performance requirements (as for the Ring), the configuration constraint may be relaxed without much penalty by moving the anode (or cold shoe) heat a short distance with heat pipes before it is radiated away. We saw earlier a number of reasons why heat pipes are favored for cooling space reactor cores. Here we review briefly how they work, to explain their utility for general transport and rejection as well. Unlike pumped fluid (single-phase) loops, which depend on the thermal capacity of the working fluid and

the temperature drop which can be established in it, heat pipes are passive, sealed, two-phase units. The working fluid vaporizes at the hot end, diffuses throughout the pipe length, condenses at the cold end, and is pulled by surface tension forces along a capillary wick back to the hot end. Using the latent heat exchanged by the fluid phase transition makes heat pipes much more efficient (as we saw quantitatively for the lithium core pipes) than their reliance on a surface tension pump would indicate. Since surface tension becomes a major fluid force in microgravity, space actually enhances the homogeneous performance of most heat pipes.

A choice of many working fluids enables heat pipes to be applied throughout the range of engineering temperatures: cryogenic (using superfluid noble gases, nitrogen, oxygen, hydrocarbon or freon) [Peterson & Compagna, 87], low (using acetone, ammonia, methanol, water or chlorofluoromethanes) [NASA TB, 85w-a], middle or power-plant waste (using potassium or sodium), and high power-plant primary (using lithium) [Merrigan, 85]. The variety of wick types and materials available and being developed all the time is so extensive --- screened, sintered, etched, chemical-vapor-deposited, brushes, grooves, channels, etcetera --- that we can regard heat pipes as practical solutions for all heat transport problems over distances of a few meters. They work in various shapes and around corners, can be made flexible [Merrigan, 85], and can be manufactured as thermal diodes [NASA TB, 81] and variable resistors [NASA TB, 83f-b]. Easily incorporated into heat exchangers [Snyder & Van Hagan, 87] and honeycomb sandwich cores [NASA TB, 80w-b & 85w-b], they join well with other thermal systems (like the pumped-fluid primary coolant loop of the Transducers' thermionic cores) and structures.

The major disadvantage of heat pipes for thermal transport in space systems is that their distributed arrangement when

ganged for large flows, which makes them extremely redundant, inevitably adds extra material to the spacecraft total. Our earlier decision to accept some size and mass penalties to gain reliability and simplicity applies therefore to heat pipes as well as reactors and converters.

Fixed passive surfaces (like the converter cold shoes), heat pipe terminals, and active-loop cold plates can all be designed as radiators, and their requirements are the same. Radiators draw heat from a system (whether it is a heat pipe, structure or other component), of course, because they are colder than that system. To remain colder, they must radiate to something yet colder. Our power radiators will have a temperature between 900 and 1200 K, while datonic components radiate at about 350 K; the temperature of deep space is 4 K. Thus the three most vital features of a space radiator are its temperature, its thermal emissivity ϵ , and what it has to look at (the effective temperature of Venus is 231 K, and the effective temperature of the sun is 5770 K). Two different approaches are available, and we will use them both.

The traditional space radiator is a second-surface mirror, typically of silver (for low solar absorptivity) fronted by a transparent coating for high IR emissivity. The short wavelengths of the incoming solar spectrum reflect back out through the clear layer, while that same layer emits the long wavelengths carrying rejected thermal energy. Selective coatings degrade in space over time [Woodcock, 86], so that even improved formulations being developed [St. Clair & Slomp, 87] cannot normally be expected to provide emittances greater than twice their absorptances after decades in space. Presuming a reference power radiator temperature of 1000 K, we will specify a surface-evaporated silver base with SiO_2 coating [Hwangbo & McEver, 85]; pure quartz is, after all, one of the most impervious materials imaginable. The silver/quartz

radiator will have an end-of-life (EOL) α/ϵ of 0.2/0.8. This is fine for our use, since as Appendix A9-1 shows, the high rejection temperature of our power plant radiators means that they function well even when they look directly at the sun or Venus. Indeed, a common material anticipated for power radiators (to which we would resort at the high end of our rejection spectrum, since silver melts at 1234 K) is carbon-carbon composite, which functions to over 2200 K [Campana, 86] but of course is highly absorptive. Silver is preferable for us also because of its high electrical conductivity, since the radiator panels are also the TE cold shoes and thus part of the power collection bus.

Appendix A9-2 shows that a more subtle approach is necessary for the low-grade radiators for power conditioning and distribution equipment, and onboard datonics packages. First, their low specific rejection capacity must be compensated in general by larger area, so these radiators will be finned. A new class of radiator [NASA TB, 8706] enhances the finned design by texturing the fins with hierarchically smaller projections, to make a cascaded blackbody. (Significant improvement resulted even for a simple unoptimized copper design, which demonstrated apparent $\epsilon > 0.6$ up to 500 K).

Of course, since it absorbs quite well also, such a blackbody radiator cannot be allowed to face another source of heat, like the sun or Venus. But Appendix A9-2 goes on to show that low-temperature radiators must be view-constrained anyway; if they faced the sun, even low-absorptivity radiators would absorb as much as they gave off, and if our blackbody radiators faced Venus, their performance would be degraded by almost 75 %. For our resonator satellites in their equatorial orbit, the only general directions assured of not viewing either the planet or the sun are those toward the ecliptic poles, which

means out the starboard and port sides of (most of) the spacecraft as they race along their orbit vectors. The low-grade radiators per se need not even face in these directions, though, because a fanned sequence of aluminized membranes interposed in the radiative path can redirect it as desired [NASA TB, 82].

As indicated in Chapter 8, our chief approach to the third category of thermal management (stabilizing the active structure) will be aggressive acclimatization. The conventional solution to maintaining dimensional accuracy in optical structures is to protect them assiduously from changing thermal influences. For the planetary resonator this is neither feasible nor sensible. The laser beam is a constant heat source, as is the dark side of Venus. But the reflected brightness of Venus' dayside changes with spacecraft true anomaly, and while the solar flux is constant between terminator crossings, the spacecraft angle to the sun line is not. So although it would be possible to estimate an orbitally-averaged spacecraft temperature, such a value would be irrelevant to the real problem, which is how the continuously changing radiative environment affects transient optical performance locally.

Trying to shield all parts of the spacecraft from the cyclic heat sources in their environment would be hopelessly impractical. Since the optical performance of the resonator and subsequent component satellites must be under continual active control anyway, thermally induced distortions will be compensated just as are dynamic disturbances. Albeit probably the greatest perturbations the satellites are likely to see, thermal effects both act more slowly than the others, and become more and more accurately predictable over the system lifetime, after hundreds of thousands of identical orbits during scores of Venusian years. This leisurely periodicity

simultaneously reduces the ongoing computational burden and increases the predictive compensation of the controlling intelligence, enhancing its successful performance.

We have already further minimized the active compensation task by making the truss members (which are most exposed to the suddenly and drastically changing solar load) of a C/Mg composite which not only is dimensionally almost immune to thermal changes, but also equilibrates its temperature quickly. This reduces the absolute and differential actuation necessary to keep each member straight and properly biased as its view of the sun (and in some cases, daylit Venus) changes. But we must insure that the active members do not get so hot in sunlight as to cook their innards. Current work on semiconductor materials like SiC and cubic boron nitride (CBN) has demonstrated componentry operable at over 920 K [Weisburd, 87], and theoretical limits are as high as 1575 K [Peterson, 87]. We have also seen earlier that fiberoptic components can work in 2200 K environments; but PZ materials, the basis of our high-frequency member actuation, depole at much lower temperatures (Chapter 8) and limit how hot we can let the structure get.

We noted earlier in this chapter a coating applied to the truss members and nodes. Although many coatings with desirable properties exist [Andus, 86], we choose a tougher type less liable than polymer-based ones to degrade in the orbital environment. It is an anodized layer processed into the surface of the magnesium composite, yielding a process-controllable α/ϵ ratio [NASA TB, 80w-a]. For reference we choose $\alpha/\epsilon = 0.3/0.7$, which in general keeps the temperature of the truss members below 335 K (Appendix A9-3). Although the anodized oxide layer is dielectric, it is also somewhat porous, so we expect any differential surface charges (resulting from the orbital particle environment) to leak

through to the conductive magnesium ground beneath. In any case, the PZ actuators, vulnerable to destruction by arcing, are protected within the conductive Faraday shells of the truss members' MMC.

Appendix A9-3 goes on to estimate the highest temperature expected for the resonator mirrors themselves, which results from solar absorption on the faces of satellites in the Station 1 constellation, at 386 K. This presumes that the rear face of each segment is anodized like the active members. Temperatures through the beryllium mirror segments equilibrate quickly also, because of the metal's high thermal conductivity. By virtue of material and coating choices then, the large active structures subjected to the resonator orbit's changing radiative conditions will distribute quickly the heat they do absorb when a source comes into view, both conductively and radiatively throughout their components, and then just as quickly reject the excess radiatively upon entering shadow. Such well-moderated and repeatably consistent behavior enables the control intelligence most simply to neutralize thermal effects on optical performance, using the actuators outlined earlier.

Our final type of thermal control consists of cooling to cryogenic temperatures any sensors (such as SQUID accelerometers) requiring it. While in this feasibility analysis it would be premature to conclude that superconductivity will always require cryogenics (Chapter 11), we should outline the manner in which such service could be provided. While the ultimate sink for thermal energy (deep space) is at 4 K, locally it is possible to achieve temperatures as low as μK [Lundholm & Sherman, 80]. Less extreme temperatures below 10 K are attainable without expendables in special closed-cycle devices. One concept is a guarded (cascaded) hydrogen refrigerator which would be light,

energy-efficient, and has only valves as moving parts [NASA TB, 85w-c]. Another, demonstrated device which is smaller but more elegant is a passive ^3He adsorption refrigerator with no moving parts at all, capable of lowering a cryostat cold plate to 0.3 K, apparently without guarding coolers [NASA TB, 83f-a]. The eventual waste heat loads of such small units are of course insignificant compared to other low-grade sources in the spacecraft.

For transporting heat isothermally over long distances from any sources buried deeply within the spacecraft to more space-exposed radiators, we can use a demonstrated chemical analog of the heat pipe [SN, 86]. In a sealed, passive, vibrationless version, heat at one end reacts a mixture of gases; after diffusing to the other end (kilometers if necessary), the products are recombined into their original mixture by a catalyst, releasing their stored chemical energy as heat and then diffusing back to the other end. For regulating conductive transport throughout sensitive regions of the bus, a variety of heat "switching" techniques exists [NASA TB, 80f-b]. Finally, a spectrum of insulation technologies will allow us to isolate spacecraft components requiring it (such as the cryostats just described), both conductively and radiatively.

Attitude Control

All the satellites must be properly "rigid body" oriented before their fine-pointing and intra-craft vernier actuation means anything. The Transducer and Ring must fly essentially sun-oriented (rotating in general once per Venusian year), but the former must be capable of three-axis rotation, and the

latter of one-axis rotation, upon retargeting. The vertex stations all fly exactly planet-oriented (rotating in general once per orbit) and require trim rotations. The separate satellites of the tethered 1δ - 1ε pair in particular require additional and continual three-axis retargeting adjustment.

Truly fast but coarse attitude control is achieved for small, especially manned, spacecraft by mass expulsion couples -- typically paired thrusters. Even apart from its logistics penalty, this technique is hopelessly inappropriate for our fleet. Gentle thrusters would lack adequate control authority for quick rotations, and powerful thrusters would introduce unacceptable point-source impulse loading into these optical spacecraft.

High-performance active orientation stability in space is typically accomplished by storing angular momentum onboard; the spacecraft uses this reservoir as both a source and sink of rotational energy for its attitude maneuvers. Conventional solutions are reaction wheels and control moment gyros (CMGs). Reaction wheels spin at high rates; conservation of angular momentum causes any change in the wheel's spin rate to be taken up by opposite bus rotation about a parallel axis through its overall mass center. At least three are needed for full three-axis control. Reaction wheels feature high precision (they will stabilize the HST), but are torque-limited [Laskin & Sirlin, 86] because their moment of inertia must be so much smaller than the bus'.

CMGs operate as large gyroscopes. The rotor spins at a constant high rate, and when controlled bearings tilt the gimbal mount, conservation of angular momentum yields a cross-product bus torque. The torque is large because it takes full advantage of the rotor spin rate (unlike the differential

rate available to a reaction wheel), but contemporary CMGs are too mechanically noisy for high precision pointing [Laskin & Sirlin, 86]. Although newly developed magnetic bearings will improve their performance greatly, even CMGs have much too little control authority for our needs. The largest CMG made to date is about as big as a desk and stores only 6100 Nms of angular momentum [Dahlgren & Taylor, 84]. By contrast, to effect even modest rotation rates, our smallest fleet craft (the Transducers) must have available almost twice that much (Chapter 4).

The system we will use throughout the fleet is a machine which combines greater torque than the CMG, with even quieter and more precise operation than the reaction wheel. Invented just over a decade ago at NASA [Anderson & Groom, 75] and tested in prototype at that time [Ball, 75], it has been named the Annular Momentum-Control Device (AMCD). Taking advantage of the result from elementary mechanics that a thin ring maximizes moment of inertia per unit mass, it eliminates entirely the rotationally inefficient hub and spokes of conventional wheels. A rim which is narrow and thin compared to its radius stores a large amount of momentum by being electromagnetically suspended, positioned and accelerated to high speed by discrete active stations spaced around its circumference.

Several features of the AMCD pre-adapt it for LSS use [Anderson & Groom, 75]. First, its lack of a "middle" removes dimension as a practical limitation; AMCDs with diameters approaching 1 km have been the topic of concerted controls research [Montgomery & Johnson, 79 & 81]. Thus control authority can grow along with spacecraft dimension, to enormous scale. Also, the open center of the attitude system can then more profitably be used for other systems, like state sensors for the AMCD [Montgomery & Johnson, 79] and the rest of the bus

[Joshi, 83] which benefit by collocation with the ring center. Suspending the rim magnetically is eased by microgravity operation; no out-of-plane nor out-of-round force bias need be considered. The hard vacuum necessary for frictionless spinning is available in orbit by default as well.

Maximum rim speed is limited by material allowable stress (Appendix A9-4), but the dominating hoop stress leads naturally to a unidirectional fiber composite layup, which takes full advantage of the fiber strength. A thin rim, of course, easily accommodates such a layup. Although the detailed AMCD design must allow for rim radial expansion from the stationary to the spun state, the hoop stress during operation stiffens the ring greatly, easing the dynamic balance task. It can be proved [Anderson & Groom, 75] that resonant coupling is impossible between the spin frequency and in-plane vibration frequencies (because the latter are always at least twice the former).

One of the greatest benefits accruing from an AMCD is its bearingless operation. Not only are wear and lubrication, and their attendant reliability complications, avoided altogether, but the magnetic suspension can prevent any vibration forcing of the bus [Anderson & Groom, 75]. Any magnetic suspension is of course a "soft mount" which automatically attenuates the high-frequency excitation that a direct mechanical connection would transmit [Laskin & Sirlin, 86]. But the sensors of a non-permanent (electromagnetic) suspension can in addition be notch-filtered at the harmonic frequencies of the spinning ring, so that the bus remains unaware of them. Since attitude control system jitter is typically a major source of vibration, the AMCD represents a qualitative advancement in spacecraft control.

Although the AMCD can be used for highly damping passive stabilization [Anderson & Groom, 75], for our mission we will

consider here only actively controlled uses. Two methods of achieving three-axis control with just one assembly are practical. The first mounts the AMCD system in a gimballed frame, which can be slewed with noncontacting magnetic bearings like a CMG. A roll (performed by changing the rim speed) about an "azimuthal" axis is followed by torquing the gimbal mount to attain the desired "elevation". After this coarse maneuver, the gimbal is locked, and torquing the rim in its magnetic gap achieves fine pointing.

The second method is yet more clever and well-suited to our needs. Two identical rims counter-rotate in a dual AMCD; setting a differential speed bias results in a roll maneuver. However, with both rims moving at the same speed, the system has no net angular momentum (Figure 9-10). An externally applied torque will cause the pair to tumble together as though they were not rotating at all. But if the two are now oppositely tipped within their magnetic gaps, momentum conservation yields a small, net "control" momentum (which because of the large rim inertia can still be quite large, as quantified in Chapter 4) to torque the spacecraft bus. Untipping the rims stops the maneuver with the craft oriented in a new position. Rapid re-orientation is possible with little bus disturbance by electronically optimizing the ramped gap tipping forces. Pointing precision, once again, becomes a function of the signal noise level rather than the mechanical system itself, despite the enormous size of the spacecraft. By specifying dual-rim AMCDs for all fleet craft we can easily achieve bus pointing accuracy well in excess of the current ST standard, strictly necessary if the final-stage EM mirror control is to perform as needed.

Ball [75] developed a scheme, based on their prototype AMCD, for a larger model suited to the active control of a Large Space Telescope. Although our AMCDs are yet much larger

in diameter, we base a reference configuration loosely on that system. First, the AMCDs "should be rings as large as the spacecraft can accommodate" [Montgomery & Johnson, 79]. The rim material must be non-magnetic and non-conducting to minimize flux drag, and have high specific working stress for mechanical efficiency. Thus we choose a Kevlar rim, with rectangular section of cm-scale [Montgomery & Johnson, 81] dimensions. Bonded into its inner edge at 10 cm intervals are (high coercivity) neo-iron permanent motor magnets, flux-linked by a ferrite band embedded beneath them. Continuous (low reluctance) ferrite bands are bonded into the "top" and "bottom" edges of the rim to complete circumferentially the flux loops of the axial suspensor fields.

Since even the linear flexibility of a large thin rim produces large absolute displacements, a tradeoff apparently exists between the out-of-plane deflections permissible and the number of discrete suspensor/drive stations desirable. In the quest for low system mass and control simplicity, the km-scale AMCDs postulated generally assume few stations, and m-scale out-of-plane excursions between them [Montgomery & Johnson, 81]. However, to control those vibration modes carefully and hence guarantee system accuracy, we will choose the opposite approach of many stations distributed around the circumference, accepting the associated power and mass penalties to buy performance. In fact, limiting the dynamic "floppiness" of the rims permits a smaller chase, which actually reduces the mass penalty of all its associated systems.

The suspensor/drive stations consist of three magnet systems: radial suspension, axial suspension, and drive stator. Electromagnets are used for all three, to enable total vibration isolation as described above. The radial suspension, acting through the ferrite band underlying the rim's permanent magnets, keeps the rim concentric with the AMCD assembly. The

axial suspension damps out-of-plane vibration modes, and applies the control torques which tip the rim. The drive stator accelerates and decelerates the rim, and supplies makeup energy for that dissipated by flux drag as the rim rotates. We specify a standard axial magnetic gap of 1 cm, a conservative (contemporary) value [Laskin & Sirlin, 86]. Motor commutation occurs by Hall sensors detecting passage of the neo-iron rim magnets. Rim axial position is monitored accurately by optical sensors.

We add one special refinement to the "standard" system package outlined above. The tremendous attitude accuracy attainable with AMCDs depends almost utterly on precise gap tipping, which as noted depends in turn on axial suspensor finesse. But for a huge rim, it depends even more basically on the relative location stability of the suspensor stations themselves. The published AMCD designs presume a rigid mounting frame, something achievable only actively for our large craft. To decouple the AMCD geometry control from the payload figure control, we need a fine-tuning actuator stage between them. Also, as a large (850 m radius) rim is spun up from rest, we can expect it to grow radially as much as 1 m; clearly the control stations must be able to move outward with it. Finally, to achieve reasonable torque authority for the largest satellites, we need an effective tipping stroke much larger than the ± 1 cm magnetic limit. We accomplish all three goals simultaneously by mounting the suspensor/drive station elements with linear PZ motors. Working together with the active bus structure, they keep the rim gap true; working cooperatively with each other, they permit a large, design-limited "effective" magnetic gap, and consequently an enhanced available control momentum.

Work on stable (robust) adaptive control for both large and multiple AMCDs has begun [Montgomery & Johnson, 79] [Joshi,

81]. Control complexity is simplified somewhat by the AMCD's inherent collocation [Joshi, 83]; its own sensors and actuators, being part of the discrete suspensor/drive stations, are close together compared to the ring diameter. The original device had two basic axial suspension modes: active positioning, and a clever stabilizing state in which "WAHOO" operation (with each suspensor having access only to state information about all the other suspensors) consumed "virtually zero power" [Ball, 75]. AMCD control must be integrated with the rest of the robot fleet intelligence.

All momentum storage devices are subject to saturation, a state in which so much accumulating momentum has been absorbed over time that the system capacity is full. We do not expect roll saturation, because the rims rotate so fast, and in opposite directions, that relative differential speed biases for roll control should not grow excessively over the AMCD lifetime. That is, none of the craft is exposed to accumulating roll torques. Some are, however, exposed to constant pitch and yaw torques (here, as in the rest of this study, the term roll means motion about the axis normal to the spacecraft bus plane, and pitch and yaw mean tilting that plane, no matter what the particular bus orientation on orbit is), notably Stations 1 α , 1 β and 1 γ .

Even AMCDs could not possibly compensate for these torques, first of all because they are huge and more essentially because they are constant. The AMCDs exist for the purpose of trim maneuvers about some nominal, "permanent" attitude setting. As outlined in Chapter 8, that constant bias is achieved through passive structural means, including both mass management and tether attachment, not by continually feeding expensive energy and expendables into the system. The tether-managed CM offset, together with propulsive couples from the ganged ion engines (described in the next section) can

provide any amount of AMCD desaturation necessary for those satellites requiring it. We know also from Appendix A8-2 that magnetic torquing, a commonly used momentum desaturation technique for low Earth orbits, should not be counted on at Venus.

Propulsion

Our choice of fleet propulsion technology is just as constrained as was our choice of power source. Assembled on-site, our large craft have no need of propulsion save for station-keeping. For the large-area resonator satellites, compensating solar pressure and solar gravity effects takes the greatest propulsive effort. For the smaller satellites of Station 1, gravitational effects from the other nearby massive craft dominate the compensation budget. For L1 and L2 stations, the station-keeping Δv itself is tiny but real, arbitrarily minimized by how close to the unstable equilibrium points they can be emplaced [Farquhar, 87]). Additionally, the Transducer must move slightly during normal retargeting maneuvers.

We want propulsive compensation to be continuous for two separate reasons. First, limiting the relative excursions of the fleet craft (by preventing displacements from growing before they are cancelled) directly minimizes the intercraft repointing, and hence processing penalty, required during operation. Second, spreading the restoring impulse more evenly over time reduces the force needed, and therefore the acceleration experienced by the satellite; as always, we seek to limit the magnitude of disturbing accelerations for these optical craft. Since propulsion by definition means exchanging

momentum with the environment, accomplishing it in orbit necessitates expelling mass; to reduce the amount of mass (logistics cost) required, we want therefore to maximize the velocity with which it is expelled, measured by the specific impulse I_{sp} .

Every one of these specifications points inevitably to electric propulsion. Generally propulsion systems (typically chemical) which can deliver high thrust suffer from low I_{sp} , while those (typically electric) which perform with high I_{sp} cannot produce much thrust. Development of new exogenous (energy supplied to the propellant from an outside source) concepts to bridge these two performance regions has been underway for some time [Finke, 80]; although many are fascinating, they are unnecessary for our fleet. Electric engines (being exogenous also) consume a huge amount of energy, but our fleet must be powered by the most energy-dense source available anyway, so electric propulsion is appropriate on all counts.

The three types of electric engines [Finke, 80] are electrothermal, electromagnetic and electrostatic. Electrothermal rockets heat a propellant to high temperatures electrically. Discharging through a gas makes an arcjet, while resistively heating a gas makes a resistojet. The latter can use practically anything (like "waste" water) for propellant, and will be used by NASA's Space Station. Both kinds expand the gas through a nozzle much like conventional chemical rockets. Electromagnetic (plasma or MPD) thrusters first ionize the propellant gas, then pass a current through it within a magnetic field. The resulting Lorentz force accelerates the ions and electrons magnetoplasmadynamically out the back, in a charge-neutralized thrust plume. A potentially higher-thrust system, this would be most appropriate for manned inter-orbit missions [Aston, 87a].

Electrostatic engines (ion engines as in Figure 9-11 are the most common) also ionize their propellant, but then accelerate (only) the atoms through a screen across a high potential difference. A separate electron gun maintains a net neutral charge on the module and hence on the spacecraft. These devices yield the highest I_{sp} of any propulsion system yet known. Ion engine technology, while still unflown, has been developed over the last quarter-century by NASA and prototypes have demonstrated reliable performance continuously for years. Current work focuses on increasing individual engine operating power (for major transportation missions) and incorporating new technology available in the coming decade [Aston, 87b].

Performance at $I_{sp} = 4250$ s, producing 0.27 N thrust at 80 % efficiency, is anticipated within a decade [DiStefano et al, 87]. Past ion engines used mercury or cesium because of their high atomic mass and low ionization potential, but now nitrogen, argon, neon, krypton or xenon are used as propellants. They simplify power conditioning, reduce engine start time, and of course are not toxic [Aston, 87b]. Furthermore, once expelled and electrically neutralized they constitute inert contaminants in the orbital environment, unlike practically all other propulsion reactants. Xenon is preferred, but because of its high atomic number it is rarer than the other gases and hence less available for large- Δv (inter-orbit) transportation missions [Pipes, 80]. However, given the cost context of the planetary laser, xenon represents the best choice for its propulsive needs.

Ion engines can be thoroughly optimized for their intended mission [Finke, 80]. In design they are scalable; in operation they gang easily, and can be throttled over a wide thrust range within their low limits. That means the small impulse they

deliver to the spacecraft can be exquisitely controlled, an advantage for our use. Our fleet will employ xenon engines in a variety of sizes, but all will actually be simpler than the devices operating even now at JPL and NASA Lewis, because they need not be gimballed. All the thrust vectoring required by our satellites is easily achieved both through their attitude control, and by applying couples using the widely-separated and throttlable ion engines themselves. Thus they can be rigidly mounted to the structural buses.

Maintenance

The vastly complex planetary laser fleet will start to die even as it is being assembled on station, a process which will continue slowly ad infinitum.

Mirror surfaces will degrade with time. Mirror EMTs will fail; each represents a single-point failure for that mirror's actuation, although the laser can function even after losing perhaps 1 % (about 14,000) of its individual mirror segments. Sensor fibers, solid state microlasers and photodiodes will break; connections will come apart. Fatigue and creep will change the tolerances of metal parts. PZ linings in active members will delaminate, crippling their operation. Arc discharges and wear will pit PZ linear motors, roughening their travel. AMCD suspensor stations will lose accuracy; rims will fray. Connector insulators will outgas, crack, shift and confuse optical and electronic telemetry. Refrigerators and heat pipes will lose their fluids. Propellant tanks will empty, and ion engines will burn out. Reactor cores will become depleted and their vessels and coolants radioactive. Converters will burn up. Micrometeoroids, charged particles

and radiation will chip away at and alter thermal control coatings. Larger impacts will sever nerve trunks, damage members and even destroy whole active neighborhoods. Modulators will wear out by losing their precision. Inter-craft telemetry lasers and star trackers will fail.

Both nominal operation and the generally benign environment of near-Venus space will eventually degrade every material system in the fleet. And cosmic rays, if nothing else, will erode its non-material intelligence as well. The subtle suck of entropy will never stop tearing apart such a finely tuned machine.

Countering such morbidity are two ongoing mechanisms. First is the control intelligence itself. Like an organic brain, its ability to carve new and redundant processing connections from its nerve networks saves it from a net loss of skill, finesse and adaptability even as individual cells die. As time goes on, it learns the performance standards and responses of its spacecraft bodies so well that it can compensate for their material degradation, certainly keeping the laser link operation within design limits until the second, repair, mechanism can help.

That help is material, from a subfleet of maintenance robots which patrol the fleet craft continuously. Maneuvering with cold nitrogen jets to preclude contamination, these itinerant bits of the fleet intelligence diagnose problems requiring close in situ inspection, reporting back their findings and mission telemetry to the larger controller. Bringing new parts from transfer craft, they repair faulty units quickly by replacement. Some robots are strong and protected enough to change out reactors, while others are dextrous enough to rig fiberoptic nerves. The most common tasks are replacing active structural members and actuators,

replenishing consumables and maintaining coatings. Like cleaner wrasses, these small robots fuss about the fleet craft, keeping them healthy on-line. Assuring accessibility to their ministrations is a major driver of detailed systems configuration.

People are in general forbidden to approach the fleet craft. The laser itself is not particularly dangerous --- even the concentrated intermediate beam, when close to Station 1, has a power density (2.3 kW/m^2) only comparable to sunlight in Venus space. But people are extremely dangerous to the laser system. Conventional biologically controlled motions are too imprecise to be tolerated near the fragile active craft, which are designed to withstand only extremely gentle or distributed disturbances. It is difficult to imagine what might need to be done to the operating satellites that the fleet brain could not figure out, and that its fleet manipulators could not execute. All human activity, being inherently space-contaminating due to debris and leakage, should be restricted near Venus.

Since the heart of each satellite is a pair of rapidly counter-spinning AMCD rims, their longevity limits each craft's lifespan. Individual suspensor/drive stations can probably be replaced on-line, but the rims themselves, being monolithic and about the same size as the whole satellite built onto them, cannot realistically be replaced without dismantling the bus structure. As their unidirectional composite layup inevitably starts to fray, it may be possible to decelerate them to a stop, effect local "patch" repairs, and then run them at somewhat reduced speeds (resulting in a graceful aging of attitude control authority), but they must be replaced well before outright failure. Should an AMCD rim unravel or split up while running at high speed, the tremendous energy stored in it would completely and explosively shred the insubstantial framework enclosing it. Thus while most of the laser system's

parts will be replaced and even upgraded individually and unobtrusively as time goes on, overhauling a satellite's AMCD rims constitutes an off-line reincarnation.

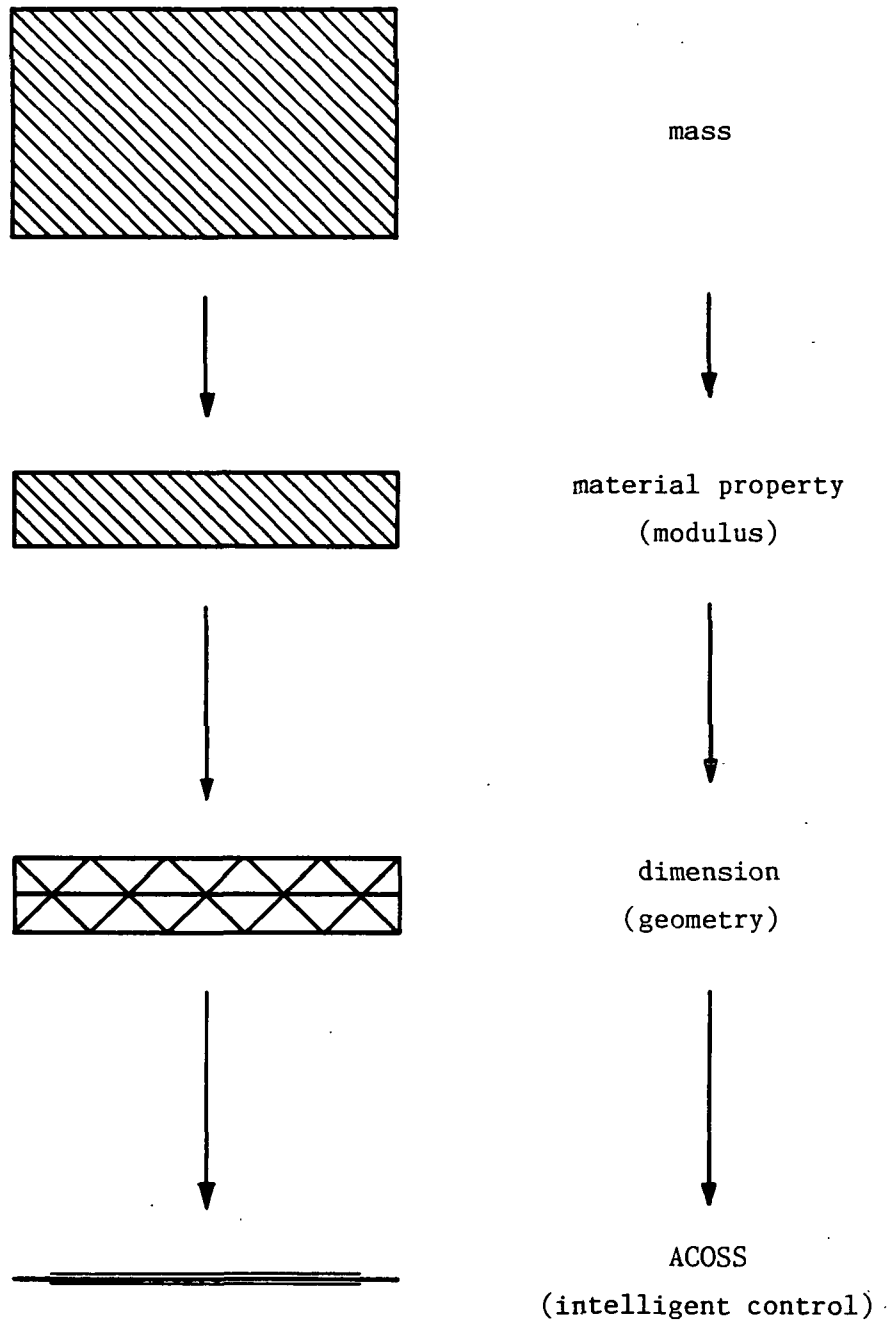


Figure 9-1 Evolution of stiffness control.

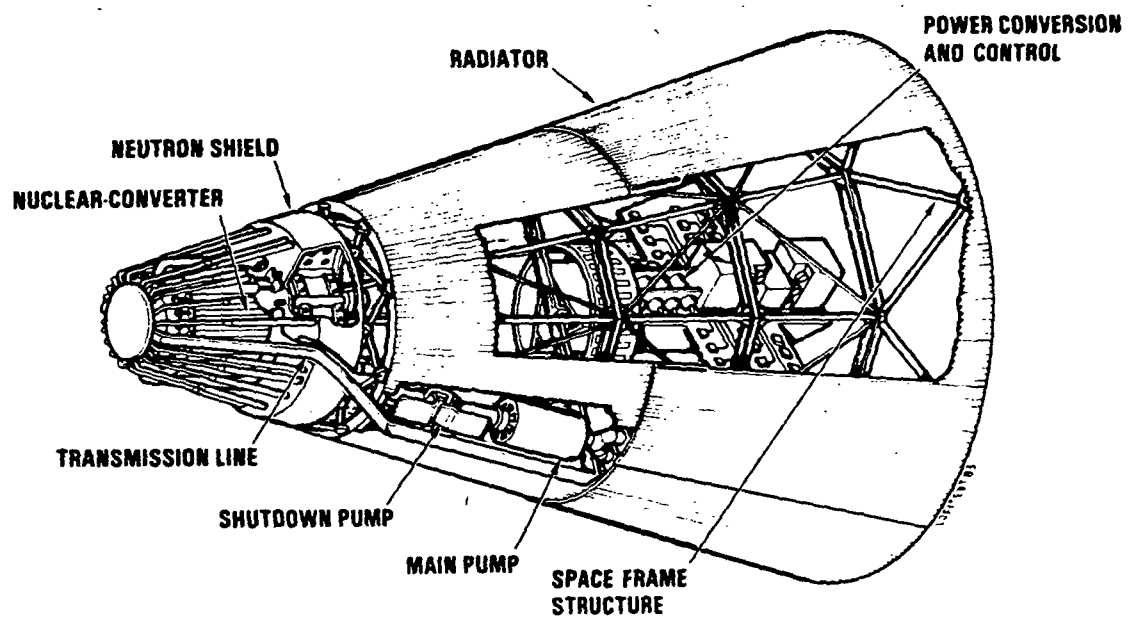
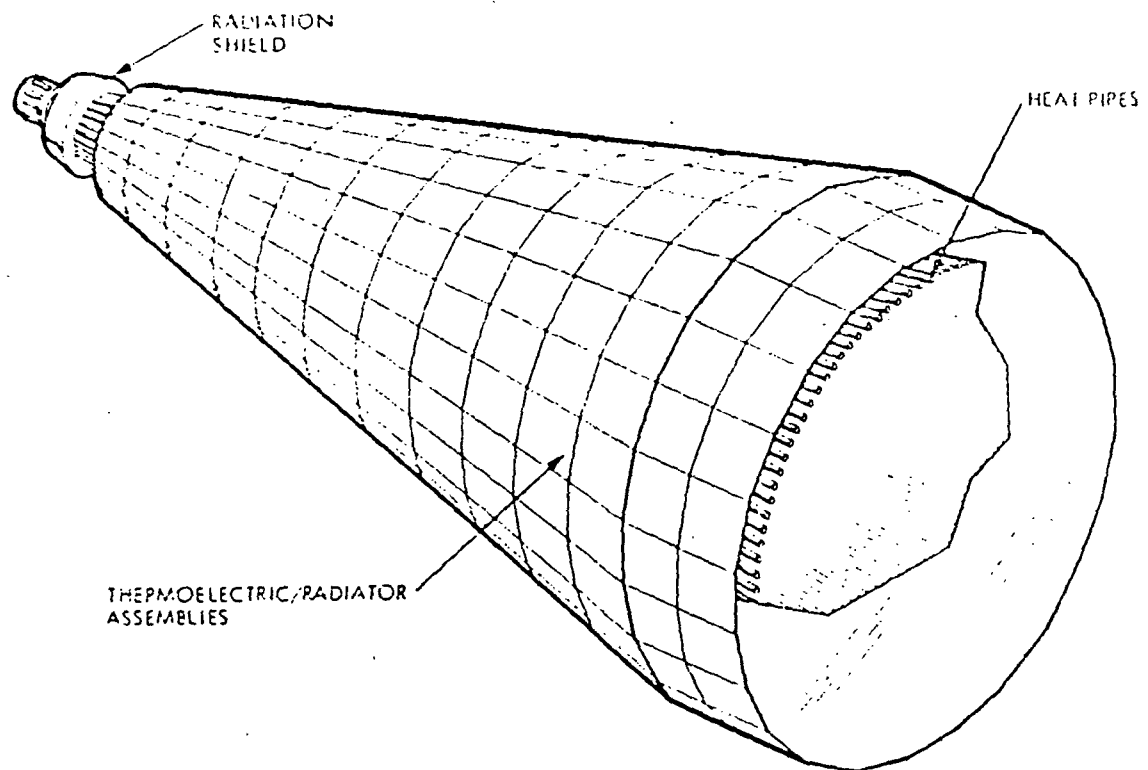


Figure 9-2 The geometrically-constrained SP-100 reactor, in its thermoelectric (above) and thermionic versions. [El-Genk & Hoover, 85]

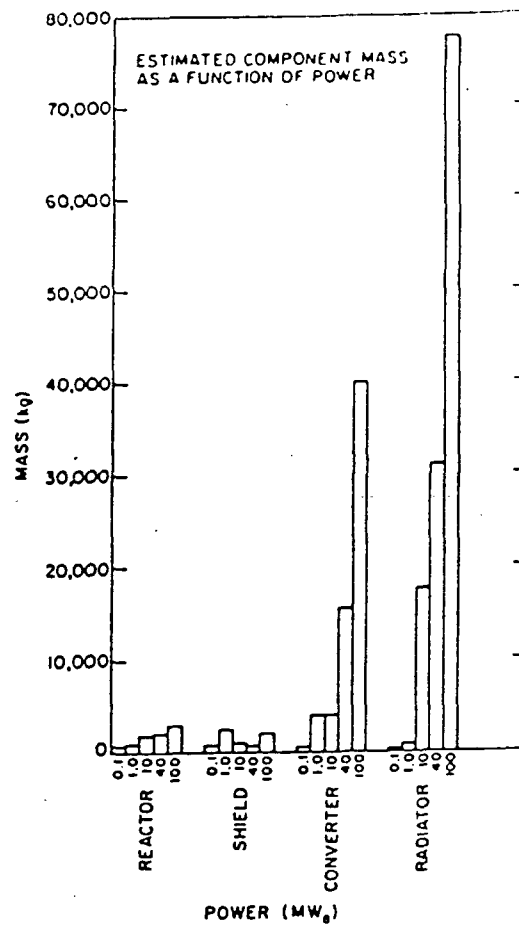


Figure 9-3 Relative sensitivity of power plant component mass to system power rating. [Feig, 85]

ORIGINAL PAGE IS
OF POOR QUALITY

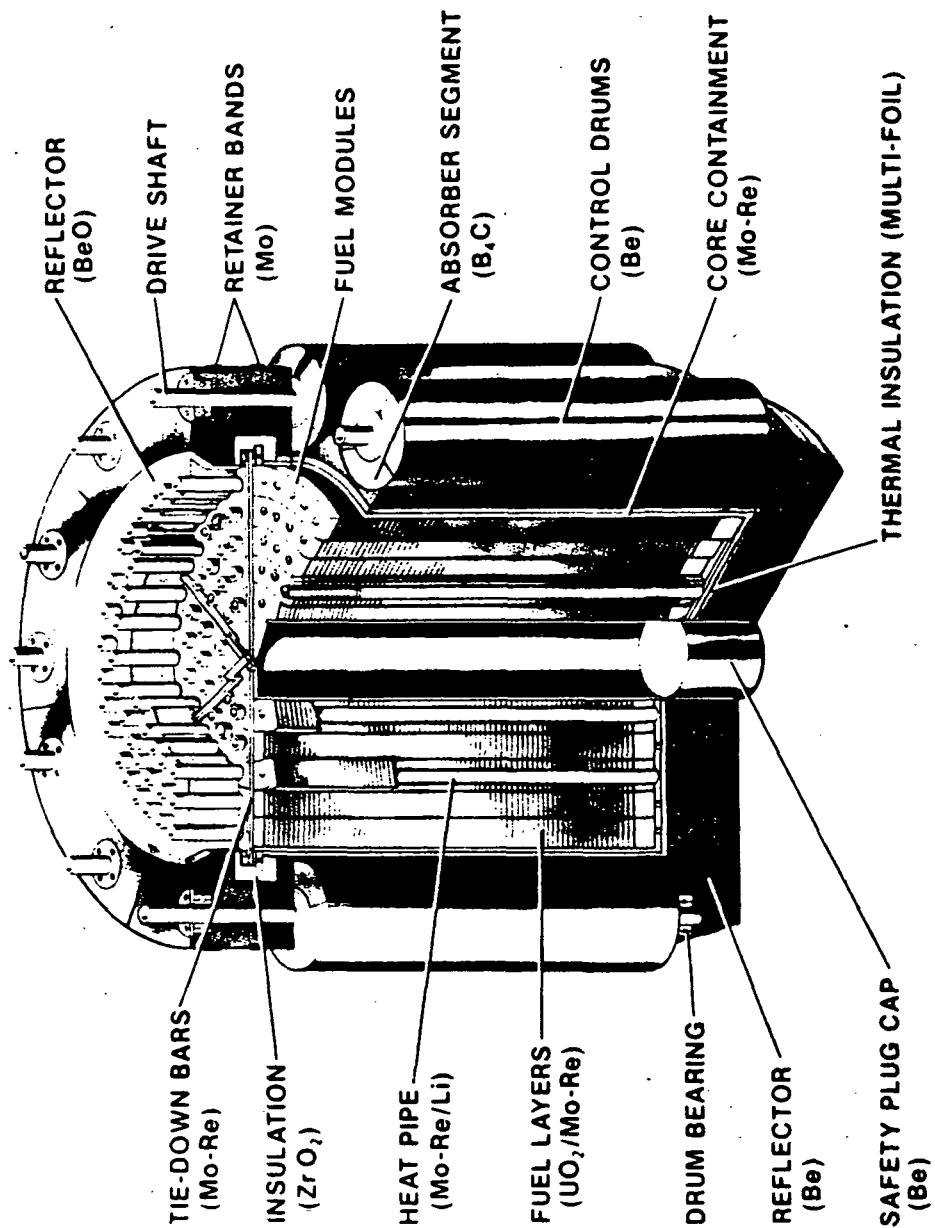


Figure 9-4 Generic heat-pipe reactor core. [El-Genk et al, 85]

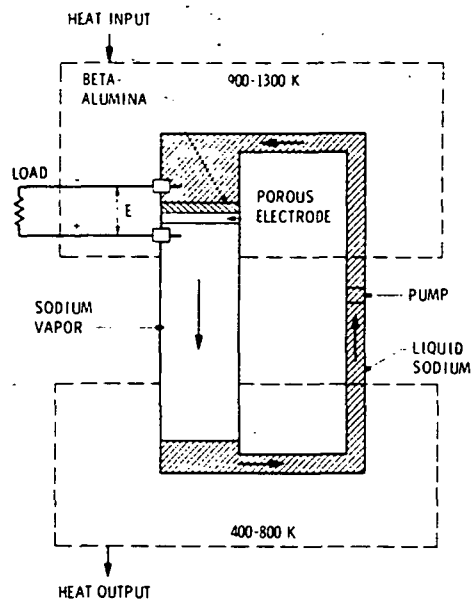


Figure 9-5 AMTEC power conversion method. [Ewell & Mondt, 85]

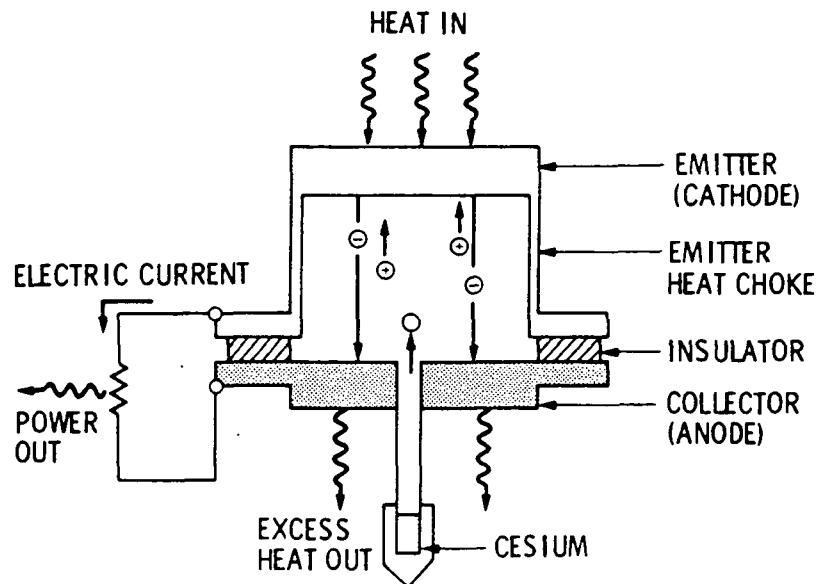


Figure 9-6 Thermionic (THI) power conversion. [Ewell & Mondt, 85]

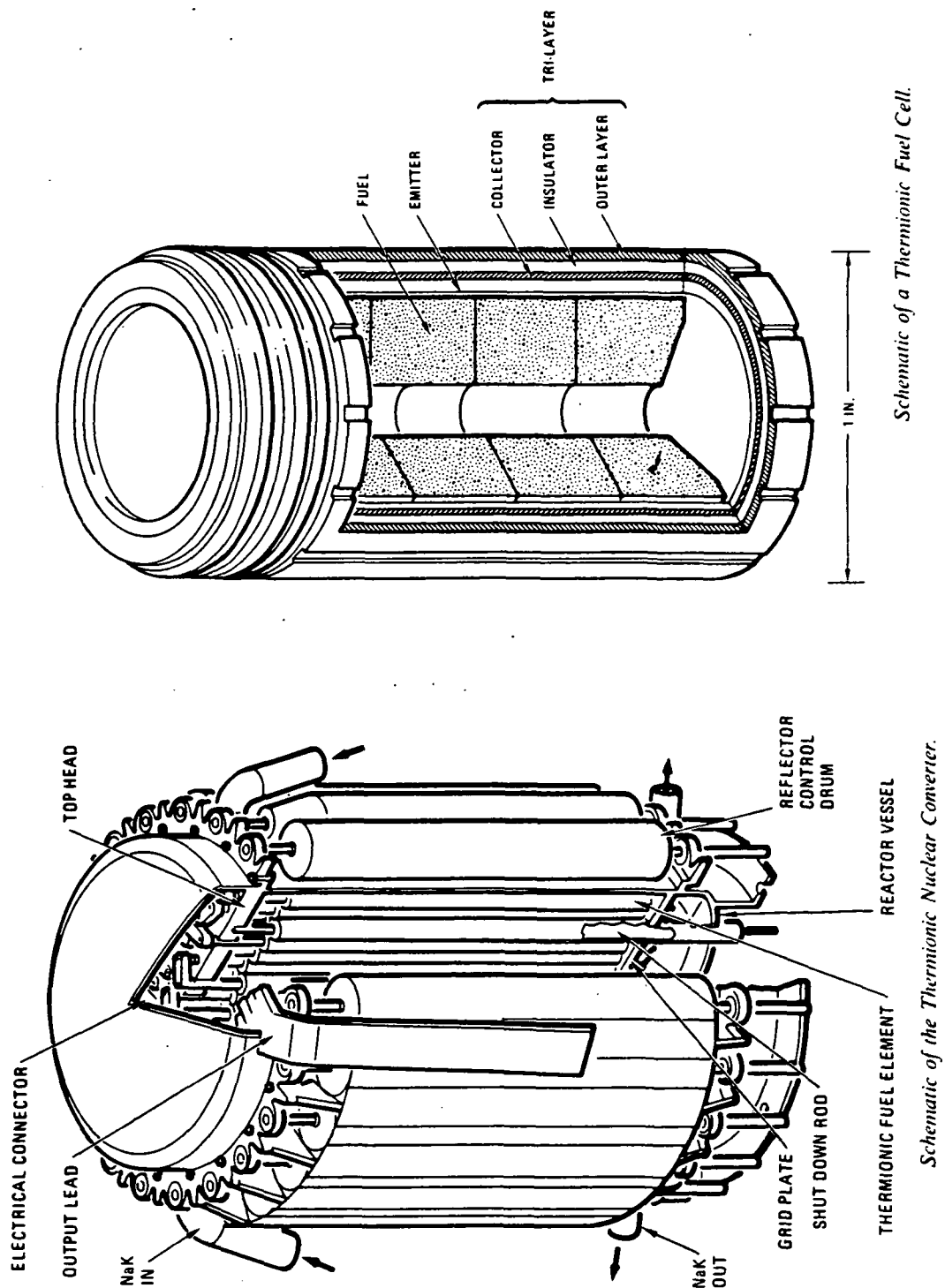


Figure 9-7 Diagram of in-core THI reactor core and a single converter cell from one Fuel Element (TFE).
[Strohmayer & Van Hagan, 85]

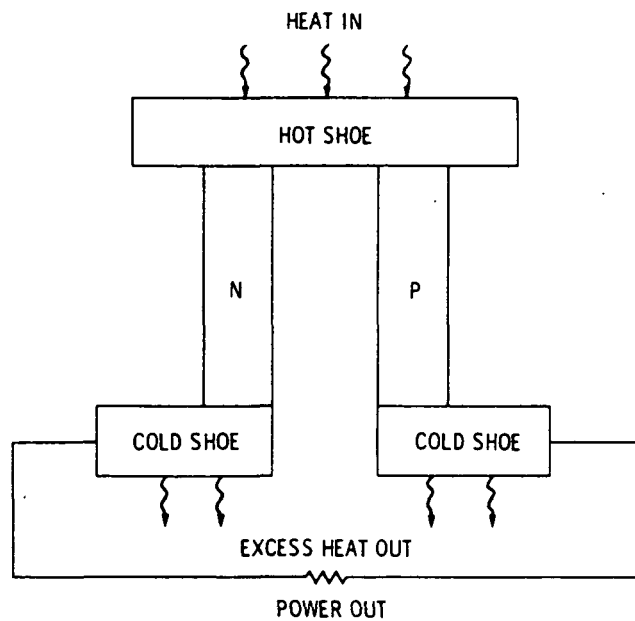


Figure 9-8 Thermoelectric (TE) power conversion. [Ewell & Mondt, 85]

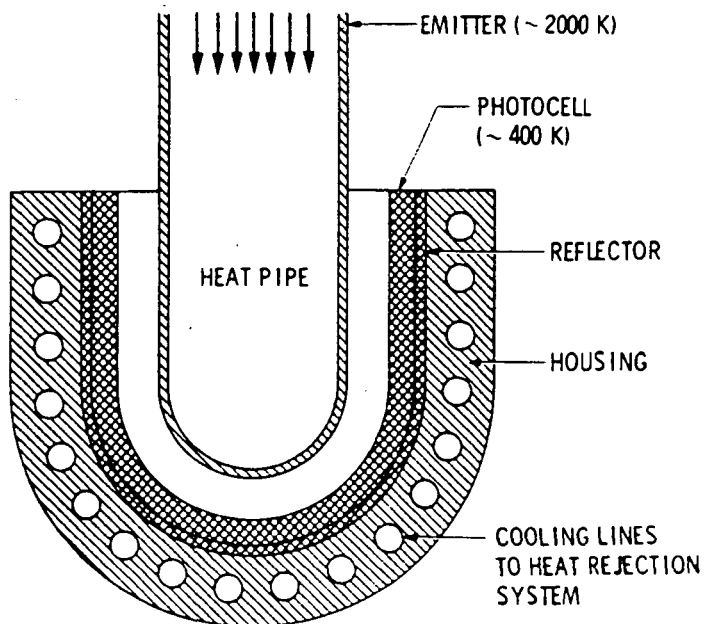
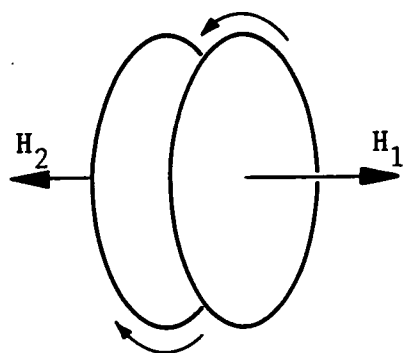
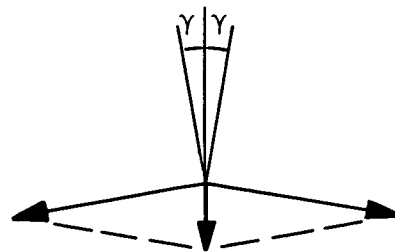
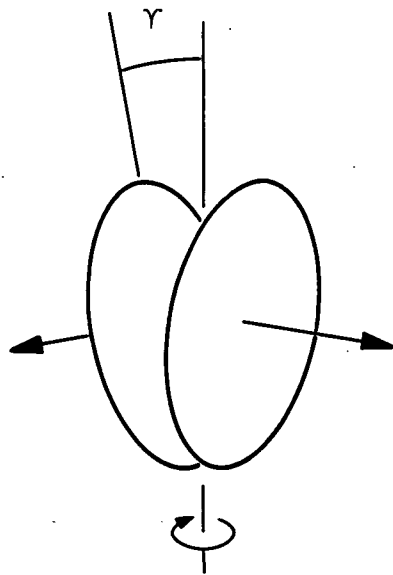


Figure 9-9 Thermophotovoltaic (TPV) power conversion. [Ewell & Mondt, 85]



$$\sum H = 0$$

System Stationary



$$\sum H = H_c$$

$$H_c = 2H \sin \gamma$$

System Rotates

Figure 9-10 Conservation of angular momentum in the AMCD.

ORIGINAL PAGE IS
OF POOR QUALITY

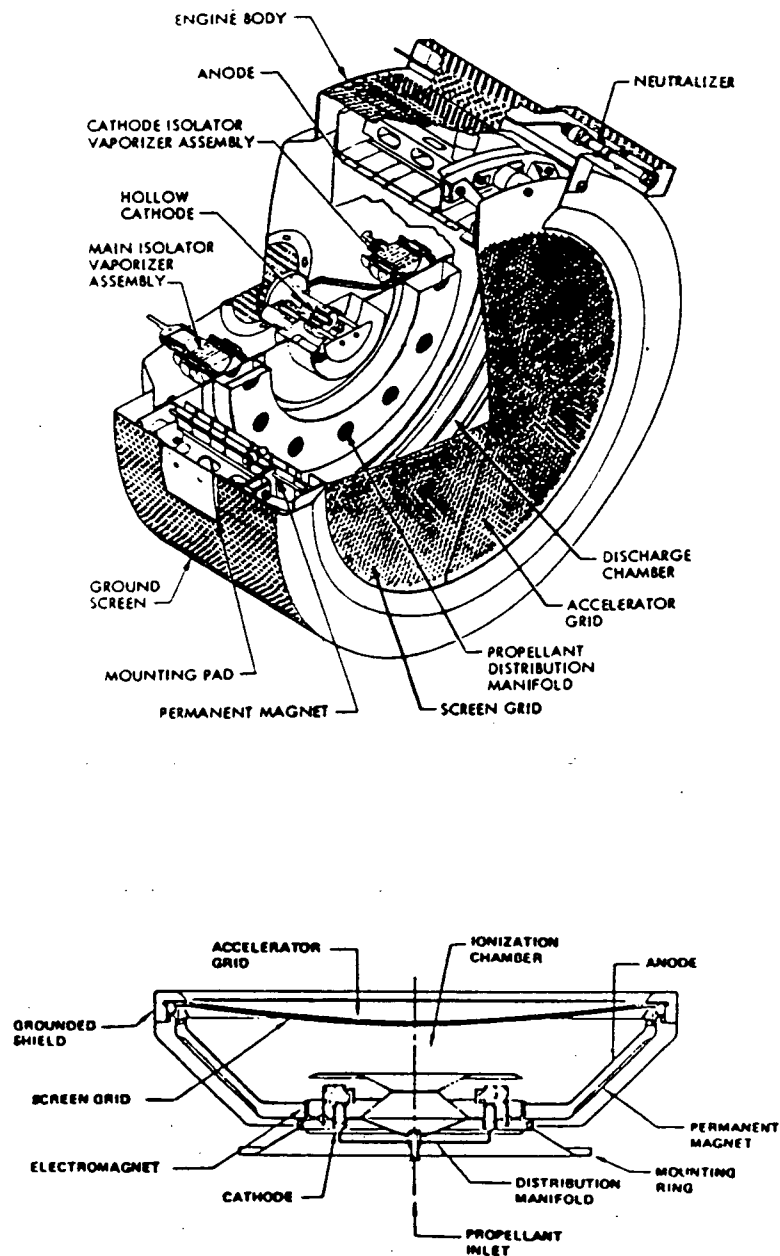


Figure 9-11 Diagrams of ion engines. [Distefano et al, 87] and [Crum, 80]

Appendix A9-1 Power radiator performance.

Viewing Space

In general the power radiators of the resonator satellites are 10 m cylinders spaced on 90 m centers. Thus their space view factor is:

$$\frac{2\pi 90 - 2(10)}{2\pi 90} = .96$$

Substituting this view factor and our assumed values for radiator emissivity and temperature into the Stefan-Boltzmann law gives:

$$\begin{aligned}\frac{\dot{Q}}{A} &= \sigma \epsilon_s F (T^4 - T_s^4) = 5.67(10^{-8})(.8)(.96)(1000^4 - 4^4) \\ &= 43.5 \frac{\text{kW}}{\text{m}^2}\end{aligned}$$

where $\sigma \equiv$ Stefan-Boltzmann constant, $F \equiv$ view factor, and we have taken the temperature of space as 4 K. Under these conditions then, the power plant radiators can reject 43.5 kW/m² to space.

Viewing the Sun

Fortunately, the sun represents a small angular source despite its high effective temperature. Presuming even a normal incidence angle, the radiators will absorb:

$$\frac{\dot{Q}}{A} = \alpha_s I_{*} = (.2) 2.6(10^3) = 0.52 \frac{\text{kW}}{\text{m}^2} \text{ of projected area}$$

or about 1.2 % of what they reject.

Viewing Venus

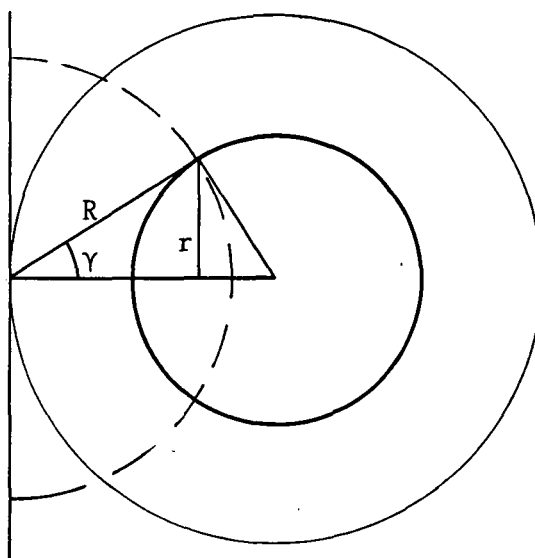
Reflected sunlight from Venus' cloud top is even less of a problem than direct solar heating, because the radiative intensity is then reduced by Venus' albedo.

Venus itself represents a thermal source with an effective temperature of 231 K on both day and night hemispheres. To ascertain the quantitative effect, we must first know what fraction of a planet-facing view is filled by the planet from our orbital altitude.

The distance R from the spacecraft to the planetary limb is:

$$(7641^2 - 6110^2)^{\frac{1}{2}}$$

or 4588 km, where 6110 km is the planetary radius out to the top of the opaque cloud layer, and 7641 km is the orbital radius. The effective radius r of the occulting disk is then:



$$r = 4588 \sin \gamma = 4588 \left(\frac{6110}{7641} \right) = 3669 \text{ km}$$

and the hemispheric view factor F_{ϕ} is simply:

$$\frac{r^2}{2R^2} = \frac{3669^2}{(2)4588^2} = .32$$

which means that Venus fills roughly a third of the sky for radiators which look in its direction. Thus the Stefan-Boltzmann law must include contributions from viewing both Venus and free space:

$$\begin{aligned}
\frac{\dot{Q}}{A} &= \sigma \epsilon_s \left[\epsilon_p F_p (T^4 - T_p^4) + (1 - F_p)(T^4 - T_s^4) \right] \\
&= 5.67(10^{-8})(.8) \left[(1)(.32)(1000^4 - 231^4) + (.68)(1000^4 - 4^4) \right] \\
&= 45 \frac{\text{kW}}{\text{m}^2}
\end{aligned}$$

is rejected from those portions of the radiators not viewing each other. The IR emissivity of Venus ϵ_p is taken as about unity, and again the temperature values are our assumed spacecraft and astronomical numbers from the previous section. This result shows directly the importance of keeping power plant rejection temperature as high as possible.

Appendix A9-2 Low-grade radiator performance.

Viewing Space

The distinguishing feature of low-grade radiation, of course, is its much lower emissive temperature. Positing an overall emissivity of 0.65 for a cascaded blackbody radiator, then even with unity view factor to space, only:

$$\begin{aligned}\frac{\dot{Q}}{A} &= \sigma \epsilon_s (T^4 - T_s^4) = 5.67(10^{-8})(.65)(350^4 - 4^4) \\ &= 553 \frac{W}{m^2}\end{aligned}$$

can be rejected. Since we already know from Appendix A9-1 that 520 W/m^2 is absorbed by a sun-facing radiator having absorptivity even as low as 0.2, clearly a low-grade radiator cannot be permitted to see the sun if it is to function.

Viewing Venus

Using the radiator emissivity and temperature cited above, and the viewing assumptions detailed in Appendix A9-1, a low-grade radiator facing Venus could reject:

$$\begin{aligned}\frac{\dot{Q}}{A} &= 5.67(10^{-8})(.65) \left\{ (1)(.32)(350^4 - 231^4) + (.68)(350^4 - 4^4) \right\} \\ &= 520 \frac{W}{m^2} \quad \text{but, given Venus' albedo of } 0.72, \text{ reflected}\end{aligned}$$

sunlight on the dayside would reduce this by as much as:

$$\frac{\dot{Q}}{A} = \alpha \alpha_F \epsilon I_* = (.72)(.65)(.32)2.6(10^3) = 389 \frac{W}{m^2}$$

Appendix A9-3 Estimated temperature highs for spacecraft components.

Truss Members

Any substantive thermal analysis would require detailed numerical simulation; here we can only perform the coarsest modeling. First we assume the member axis is normal to the sun line, as for example would be the case for a member in the top layer of the mirror support structure when the resonator satellite crossed the subsolar point. We assume the entire spacecraft bus to be in thermal equilibrium, so that no heat is exchanged through the cylindrical half of the member which faces "into" the spacecraft. We thus treat only the "outer" half, which absorbs solar heat in proportion to its projected area (diameter) but radiates heat to space in proportion to its half-circumference. Considering then only solar heat loading (ignoring the member's own internally generated heat from datonics, PZ and thermal actuators), equilibrium requires that the member temperature be:

$$T_{\text{equil}} = \left[\frac{2\alpha I_{\text{sun}}}{\pi \sigma \epsilon F} + T_s^4 \right]^{\frac{1}{4}} = \left[\frac{2(.3)2600}{5.67(10^{-8})(.7)(1)} + 4^4 \right]^{\frac{1}{4}}$$
$$= 334 \text{ K}$$

where the variables are defined as in Appendices A9-1 and A9-2, and we use $\alpha/\epsilon = 0.3/0.7$ for the anodized thermal control coating on the active structural member. The radiative view factor to space is taken as unity for this outer half of the member.

Mirror Segments

Obviously the sun cannot shine on both the truss members and the mirror faces at the same time, since they are on opposite sides of the

practically two-dimensional satellite. The gold surface of the beryllium mirrors determines their solar properties; total spectral solar absorptivity may be taken as $\alpha = 0.17$ [Berggren & Lenertz, 75]. The mirror backs are presumed treated by thermal-control anodization like the structural members; assuming a 50 % view factor of the mirror backs through all the trusswork, we find the solar absorptivity to be $(.5)(.3) = 0.15$, comparable to the value we use for the front. Hence we may to first order conclude that the mirrors equilibrate to the same temperature regardless of whether they face into or away from the sun. In any case, we ignore as inconsequential the heat load from the laser itself, since the beam density is only 11 W/m^2 and the mirrors are designed for peak IR $\alpha < 0.005$. The mirrors absorb:

$$\frac{\dot{Q}}{A} = (.17)2600 = 442 \frac{\text{W}}{\text{m}^2}$$

Even assuming that this heat could only be radiated out the mirror back, with that aforementioned 50 % view factor, leads to a peak "equilibrium" temperature for the mirrors of:

$$T_{\text{equil}} = \left[\frac{442}{5.67(10^{-8})(.7)(.5)} + 4^4 \right]^{\frac{1}{4}} = 386 \text{ K}$$

This peak would apply, for instance, to the mirrors of 1α , which do actually face directly into and away from the sun at particular orbital epochs.

Appendix A9-4 Maximum AMCD rim speed.

For a spinning thin ring, tangential stress dominates, and the maximum rotational speed ω_{\max} is given by [Anderson & Groom, 75]:

$$\omega_{\max} = \left(\frac{\sigma}{\rho} \right)^{\frac{1}{2}} \frac{1}{r}$$

where $\rho \equiv$ the ring material density, $r \equiv$ the ring radius (assumed large compared to the ring thickness), and $\sigma \equiv$ the material working stress.

The quantity $\frac{\sigma}{\rho}$ is called the specific working stress, and its square root, measured in m/s, serves as a figure of merit for ranking materials for high-speed spinning rings. This value for aluminums is less than 350 m/s, and even for the best steels is less than 500 m/s. Unidirectional layups of boron or graphite fibers do better, at 1060 m/s, and Kevlar aramid fiber is the current best, at 1230 m/s.

Although better materials are sure to come along, we assume Kevlar AMCD rims throughout the fleet. The maximum spin rates for rims of various relevant radii are calculated as:

<u>r</u>	<u>ω_{\max}</u>	
(m)	(rad/s)	(rpm)
7.5	164	1566
100	12.3	118
650	1.89	18.1
850	1.45	13.8

References

Aerospace America- "Aerospace Materials: Weld Metal Matrix Composites" (Oct 1987).

Willard W Anderson & Nelson J Groom- The Annular Momentum Control Device (AMCD) and Potential Applications
NASA Technical Note TN D-7866 (Mar 1975).

J L Anderson & L C Oakes- "Instrumentation and Controls Evaluation for Space Nuclear Power Systems"
in [El-Genk & Hoover, 85].

Andus Corporation- "Spacecraft Radiative Temperature and Charge Control Films" product literature (Mar 1986).

David A Arnold- "The Behavior of Long Tethers in Space"
AAS 86-202 in [Bainum et al, 87].

Graeme Aston- (a) "Transportation Applications of Electric Propulsion" delivered at 25th Goddard Memorial Symposium
Visions of Tomorrow: A Focus of National Space Transportation Issues NASA Goddard Space Flight Center
(Mar, 1987).

- (b) "Ferry to the Moon" Aerospace America
Vol 25 No 6 (Jun 1987).

Sterling Bailey, Sam Vaidyanathan, John Van Hoomissen- "Liquid Metal Cooled Reactors for Space Power Applications"
in [El-Genk & Hoover, 85].

Peter M Bainum, Ivan Bekey, Luciano Guerriero, Paul A Penzo
(eds)- Tethers in Space Advances in the Astronautical Sciences Vol 62 (Univelt, 1987).

Ball Brothers Research Corporation- Annular Momentum Control Device (AMCD) Final Report NASA CR-144917 (1975).

W A Baracat & C L Butner- Tethers in Space Handbook Code MT, NASA Headquarters (NASA, 1986).

William J Barattino, Mohamed S El-Genk, Susan S Voss- "Review of Previous Shield Analysis for Space Reactors" in [El-Genk & Hoover, 85].

Franco Bevilacqua, Pietro Merlina, Alberto Anselmi- "Tethered Platforms: New Facilities for Scientific and Applied Research in Space" AAS 86-238 in [Bainum et al, 87].

David Buden & James H Lee Jr- "An Approach to Space Reactor System Selection and Design" in [El-Genk & Hoover, 85].

Delbert F Bunch- "Space Reactor Systems and Safety Strategies" in [El-Genk & Hoover, 85].

R J Campana- "Heat Radiators for Electromagnetic Pumps" JPL Invention Report NPO-16458/SC-1411 for NASA Tech Briefs Vol 10 No 5 Item 81 (Sep 1986).

Gordon L Chipman Jr- "Space Nuclear Power Programs: Present and Future" in [El-Genk & Hoover, 85].

Franklin Hadley Cocks- "Ultralight Reactive Metal Foams in Space: A Novel Concept" Journal of Spacecraft and Rockets Vol 21 No 5 (Sep-Oct 1984).

C M Cox, D S Dutt, R A Karnesky- "Fuel Systems for Compact Fast Space Reactors" in [El-Genk & Hoover, 85].

Earle M Crum- "Electric Propulsion for SPS" in [NASA CP-2144].

John B Dahlgren & Lawrence W Taylor Jr- "Spacecraft Control Research at NASA" [op cit Chapter 8].

E DiStefano, G A Beale, T J Trapp, D Bohl- "Space Nuclear Propulsion: Future Applications and Technology" in [El-Genk & Hoover, 87].

Mohamed S El-Genk & Mark D Hoover- Space Nuclear Power Systems Proceedings of 1st & 2nd Symposia on SNPS, Albuquerque Vols 1 & 2 (Orbit Book, 1985) and Vols 3 & 4 (Orbit Book, 1987).

Mohamed S El-Genk & Jong-Tae Seo- "Trends and Limits in the Upgrading of SP-100 Baseline Design of Nuclear Powered Space System" in [El-Genk & Hoover, 87].

Mohamed S El-Genk, David M Woodall, Virginia F Dean, David L Y Louie- "Review of the Design Status of the SP-100 space Nuclear Power System" in [El-Genk & Hoover, 85].

David G Elliott- "Rotary Radiators for Reduced Space Powerplant Temperatures" in [El-Genk & Hoover, 85].

R Ewell & J Mondt- "Static Conversion Systems" in [El-Genk & Hoover, 85].

Robert W Farquhar- personal communication, NASA Goddard SFC (Jun 1987).

Jason R Feig- "Radiator Concepts for High Power Systems in Space" in [El-Genk & Hoover, 85].

Robert C Finke- "Electric Propulsion Technology" in
[NASA CP-2144].

Ronald E Glickman & Samuel C Rybak- "Gravity Gradient
Enhancement During Tethered Payload Retrieval" AAS 86-223
in [Bainum et al, 87].

J W Holland, M H Horner, L Yang- "Thermionic Fuel Element
Technology Status" in [El-Genk & Hoover, 85].

W G Homeyer, M H Merrill J W Holland, C R Fisher, D T Allen-
"Thermionic Reactors for Space Nuclear Power" in [El-Genk
& Hoover, 85].

Han Hwangbo & W S McEver- "Thermal Management of High Power
Space Based Systems" in [El-Genk & Hoover, 85].

S M Joshi- "Control of Large Space Structures Using Annular
Momentum Control Devices" Spacecraft Pointing & Position
Control Advisory Group for Aerospace Research and
Development AGARD-AG-260 (Nov 1981).

- "Robust Precision Pointing Control of Large Space
Platform Payloads" Structural Dynamics and Control of
Large Space Structures, 1982 NASA Langley Research Center
NASA-CP-2266 (Apr 1983).

Serope Kalpakjian- Manufacturing Processes for Engineering
Materials (Addison-Wesley, 1985).

A Kirpich & S Rao Yadavalli- "Regulation Options for
Spacecraft Thermoelectric Power Converters"
in [El-Genk & Hoover, 87].

Daniel R Koenig- "Heat Pipe Reactor Designs for Space Power"
in [El-Genk & Hoover, 85].

R A Laskin & S W Sirlin- [op cit Chapter 8].

Larry G Lemke, J David Powell, Xiaohua He- "Attitude Control of
Tethered Spacecraft" AAS 86-211 in [Bainum et al, 87].

J G Lundholm Jr & Allan Sherman- "A Review of the NASA/OAST
Cryogenic Coolers Technology Program" Heterodyne Systems
and Technology conference proceedings, NASA Langley
Research Center (Aug 1980).

Tom Mahefkey- "Overview of Thermal Management Issues for
Advanced Military Space Nuclear Reactor Power Systems"
in [El-Genk & Hoover, 85].

Manuel Martinez-Sanchez & D E Hastings- "A Systems Study of a
100 kW Electrodynamic Tether" AAS 86-224
in [Bainum et al, 86].

A T Mattick & A Hertzberg- "Liquid Droplet Radiators for Heat
Rejection in Space" Journal of Energy Vol 5 No 6
(Nov/Dec 1981).

D L McDanel, T T Serafini, J A DiCarlo- "Polymer, Metal, and
Ceramic Matrix Composites for Advanced Aircraft Engine
Applications" Journal of Materials for Energy Systems
Vol 8 No 1 (Jun 1986).

M A Merrigan- "Heat Pipe Technology Issues" in [El-Genk &
Hoover, 85].

Military Space- "New Space Power Source?" and "DOD Highlights
Soviet Space Program" (23 Apr 1984).

Mohan Misra- "Metal-Matrix Composites" delivered at
Aerospace: Century XXI AAS 33rd Annual Meeting, Boulder
(29 Oct 1986).

Raymond C Montgomery & C Richard Johnson Jr- "The
Dual-Momentum Control Device for Large Space Systems"
AIAA paper 79-0923 (1979).

- "Adaptive
Control System for Large Annular Momentum Control Device"
The Microwave Radiometer Spacecraft: A Design Study
NASA Langley Research Center RP-1079 (Dec 1981).

NASA Conference Publication 2144- Large Space
Systems/Low-Thrust Propulsion Technology
NASA Lewis Research Center (May 1980).

NASA Tech Briefs- "Peeled-Film Solar Cells" p 135
(Summer 1980).

- "Composites with Nearly-Zero Thermal
Expansion" p 323 (Fall 1980) (a).
- "Heat Switch Has No Moving Parts" p 359
(Fall 1980) (b).
- "User Chooses Coating Properties" p 451
(Winter 1980) (a).
- "Heat Pipes Cool Probe and Sandwich Panel"
p 478 (Winter 1980) (b).
- "Orifice Blocks Heat Pipe in Reverse Mode"
p 185 (Summer 1981).

- "Reflective-Shield Radiative Cooler" p 423
(Spring/Summer 1982).
- "Passive Module for Cryogenic Refrigeration"
p 93 (Fall 1983) (a).
- "Variable-Conductance Heat-Transfer Module"
p 98 (Fall 1983) (b).
- "Thermoelectric Generator" p 347
(Spring 1984).
- "Static-Suppressing Optical Paint" p 107
(Spring 1985).
- "Titanium Heat-Pipe Wicks" p 141
(Winter 1985) (a).
- "Thermal-Diode Sandwich Panel" p 142
(Winter 1985) (b).
- "Hydrogen Refrigerator Would Cool Below
10 K" p 153 (Winter 1985) (c).
- "Solar Cells With Reduced Contact Areas"
p 26 (Jan 1987).
- "Metal-Clad Graphite/Epoxy Tubes" p 73
(May 1987).
- "Cascaded-Blackbody Heat Radiators" p 33
(Jun 1987).

Ralph F Orban- "Development, Testing, and Evaluation of New Tether Materials" AAS 86-243 in [Bainum et al, 87].

Judy K Partin- "Radiation Effects Testing of Optical Fibers in a Nuclear Reactor" in [El-Genk & Hoover, 87].

James Francis Peters- Nuclear Power in Aerospace Vehicles: A Technical Overview M.S. paper UMAERO-29 University of Maryland, College Park (1987).

G P Peterson & G L Compagna- "Review of Cryogenic Heat Pipes in Spacecraft Applications" Journal of Spacecraft and Rockets Vol 24 No 2 (Mar/Apr 1987).

Ivars Peterson- "Putting the Heat on New Semiconductors" Science News Vol 132 p 247 (17 Oct 1987).

William E Pipes- "DOD Low-Thrust Mission Studies" in [NASA CP-2144].

O Remondiere, R Pailler, A Momode, Ph. Roy- "Magnesium Matrix Composite Materials Processing and Applications" Proceedings of the First European Conference on Composite Materials and Exhibition Bordeaux France (24-27 Sep 1985).

E Scala, Lee S Marshall, Douglas P Bentley- "Design and Fabrication of the 20 km / 10 kV Electromechanical Tether for TSS-1 Using High Impact Conductor" AAS 86-248 in [Bainum et al, 87].

Science News- "Moving Heat Around, Chemically" Vol 129 No 25 (21 Jun 1986).

Stepan S Simonian- [op cit Chapter 8].

H J Snyder & T H Van Hagan- "Heat-Exchanger/Heat-Pipe Interface" JPL Invention Report NPO-16456/SC-1400 for NASA Tech Briefs Vol 11 No 2 Item 6 (Mar 1987).

J Soovere & M L Drake- Aerospace Structures Technology Damping Design Guide 3 Vols Air Force Wright Aeronautical Laboratories AFWAL-TR-84-3089 (Dec 1985).

Anne K St Clair & Wayne S Slemp- "Evaluation of Colorless Polyimide Film for Thermal Control Coating Applications" Technical Support Package LAR-13539 for NASA Tech Briefs (Apr 1987).

W H Strohmayer & T H Van Hagan- "Parametric Analysis of a Thermionic Space Nuclear Power System" in [El-Genk & Hoover, 85].

B J Webb & Z I Antoniak- "Rotating Bubble Membrane Radiator for Space Applications" Proceedings of 21st IECEC (Aug 1986).

S Weisburd- "They May Not Be Super, but Semis Are Hot" Science News Vol 132 p 390 (19 & 26 Dec 1987).

James R Wertz (ed)- [op cit Chapter 6].

W Westphal & G Kruelle- "Alternative Space Power Systems" DGLR Jahrestagung Bonn (Sep-Oct 1985).

J R Wetch, C J Nelin, E J Britt, G Klein- "Reactor Power System Deployment and Startup" in [El-Genk & Hoover, 85].

Charles Wood- "High-Temperature Thermoelectric Energy
Conversion" JPL Invention Report NPO-16548/6041 for
NASA Tech Briefs Vol 11 No 3 Item 128 (Mar 1987).

Gordon R Woodcock- [op cit Chapter 7].

PART 3

CONTEXT AND MEANING

Ahnest du den Schöpfer, Welt?

-- Friedrich Schiller

CHAPTER 10

NON-PLANETARY LASERS

Chapter Abstract - Natural planetary lasers constitute a limited resource, and available planets supporting atmospheric CO₂ lasing may be lacking in other star systems. Establishing an effective interstellar communication network to assist human expansion beyond our native solar system must depend on alternatives. Many candidates for high-power space lasers, including solar-pumped CO₂, have been proposed for many uses. Such devices could match the data transfer rates available to planetary lasers by compensating with greater beam power for wider emission spectra. Separate transmitter facilities could then be constructed to maintain dedicated continuous contact with several target star systems. Materials technology, rather than controls evolution, limits system development for non-planetary lasers.

The Need for an Alternative

We have up to this point erected an elaborate structure of spacecraft design and performance on the foundation of natural laser emission discovered at Mars and Venus. But like an inverted pyramid, that superstructure is only as solid as the point on which it rests. Reiterating then that the mission presumed by Chapter 1 (interstellar communication at high data transfer rates) is indeed important to the fate of intelligence in the galaxy, leads us for three strong reasons to investigate alternative methods for accomplishing that mission, which do not depend on planetary lasers.

First and most obvious is that the planetary laser may turn out to be unworkable after all. The natural solar-pumped emission may be intermittent, or vary too unreliably to be used for an information carrier. These possibilities can be settled only by empirical study. Or both Venus and Mars may prove unavailable because of other human projects. Mars remains a popular target for human activity incompatible with a planetary resonator; Venus may fall prey to terraforming (although detailed study indicates such a project would consume over 16 millennia and 10^{30} J [Fogg, 87]). Even if available, the Mars laser may prove unusable because of intrinsic problems discussed in Chapter 6. Finally, practically achieving the fine conditions required to operate any planetary resonator may prove even more difficult than we anticipate (Chapter 7), due to reasons either beyond the scope of this study or growing from seeds contained within it. In short, any feasibility we have established for planetary lasers depends on assumptions made along the way, all of which need more detailed study.

Second and more farsighted is that, even should it work, a planetary laser constitutes an extremely limited resource.

If we made requited contact with several alien cultures, or if we maintained contact with several distant human colonies, or both, we would inevitably desire full duty cycles to each target. The singular link allowed by a planetary laser would be a frustrating bottleneck under such circumstances. Now since a pentagonal resonator path at Venus, for instance, only intersects about 5 % of the available gain medium under its orbit, we might imagine up to 20 interleaved pentagonal resonators orbiting the planet. That would alleviate but not eliminate the ultimate system capacity problem.

Third and most central is that, while perhaps not unique, planets supporting atmospheric CO₂ lasers are probably not ubiquitous either. Assuming now a real need to keep in reciprocal contact with stellar human colonies, we must address the possibility that many eligible star systems may lack planets altogether. That is, a materially rich solar system would be most exploitable if its matter were bound in multitudes of planetesimals rather than gravitationally concentrated planets. Great hope of finding around some other star a planet supporting liquid water is not justified anyway (Chapter 1); still, it might even happen that probes would find a system having an Earthlike planet but none with CO₂ atmospheres! The settling culture could receive lots of information from the Venusian laser, but be unable to respond in the same way.

Together these three reasons mean that establishing the feasibility of practical interstellar communication must include alternative methods. This chapter thus investigates prospects for non-planetary lasers achieving the same interstellar data transfer performance as planetary lasers.

High-Power Space Lasers

The link equation (equation A3-3.7) provides a convenient tool for examining fundamental differences between planetary and non-planetary lasers. In comparing the two types, we take all properties of the link itself, including beam spreading efficiencies, target spot size control, bandwidth-matched heterodyne detection, and acceptable channel SNR, to be the same. The useful channel capacity B is then directly proportional to the square of the transmitted power P_s and inversely proportional to the source laser linewidth B_s (which determines the matched-bandwidth condition).

The reference planetary laser from Part 2 has a rather low transmitted power (180 kW, Appendix A7-13), considering its huge intracavity beam volume. But the extremely narrow emission linewidth (3.3 Hz, Appendix A7-12) enforced by its comparably huge cavity length results anyway in a high specific power (W/Hz), making it detectable with useful SNR by a narrow-channel receiver. In designing a non-planetary laser for the same job, we exploit this tradeoff in the other direction. If we assume a double-pass cavity length on the order of 1 km (instead of the 90,000 km necessitated by the Venusian pentagonal laser), the intracavity beam would need a coherence time of order only $3 \mu s$ to maintain lasing, limiting the allowable oscillation linewidth to no more than about 300 kHz (instead of the 3.3 Hz maximum enforced by Venus).

The cavity control to generate such a linewidth is much simpler than that required by the planetary resonator, and we reasonably assume based on the technology outlined in Part 2 that much narrower linewidths than 300 kHz are achievable in general for space-based lasers. (Incidentally, a

matched receiver having 300 kHz channels would need far fewer of them to track signal excursions than the one necessitated by the 3.3 Hz planetary laser beam.) But to complete the comparison, let us indeed take the transmitter spectral linewidth as 300 kHz. To achieve the same useful channel capacity as the reference planetary laser then, this one would need to put out 54.3 MW of beam power. A large but non-planetary laser could thus match the performance of the Venus laser at output powers less than of order 50 MW.

A 50 MW power ceiling conservatively fits the range envisioned by published studies of high-power space laser applications. Without even invoking SDI, for instance, centralized laser systems supplying remote spacecraft power would need to produce of order 10 MW, whereas systems for rocket propulsion or terrestrial power supply would need of order 1 GW, a rating thought to be achievable by only a decade of concerted development [Humes, 82] [Lancashire, 82a]. Several conceptual designs for 100 MW laser plants have been proposed [Holloway & Garrett, 82], and many system designs for 1 MW plants have been detailed and even costed [Lancashire, 82b] [Prelas et al, 85] [De Young et al, 87].

A variety of laser candidates supports these proposals, with wavelengths ranging typically from the visible to the IR. Pumping schemes include direct nuclear, conventional electric discharge, catalyzed chemical reactions, and direct and indirect solar [Taussig et al, 79] [Weaver & Lee, 81] [Holloway & Garrett, 82] [Williams & Zapata, 85] [Conway & De Young, 85]. Proposed lasants include solid neodymium slabs, liquid Nd^{3+} , aerosol I_2^* , metal vapors, gaseous CO , CO_2 , IBr , and the organic iodides CF_3I , $\text{C}_3\text{F}_7\text{I}$ and $\text{C}_4\text{F}_9\text{I}$. Gaseous lasants might operate in either gasdynamic or static cavities. Really, just about every kind of laser system known has been considered to some degree for high-power remote applications in space.

While recognizing that it represents a somewhat arbitrary choice from that large menu, we concentrate here specifically on solar-powered, optically-pumped CO₂ lasers producing light at 10.6 μm , because of the reference frame defined by the rest of this study, including heterodyne detection. The most well-defined scheme (indirectly pumped) is capable of lasing efficiencies as high as 20 %, leading to an assumed space system plug efficiency on the order of 10 % [Lancashire, 82c].

Indirect pumping is a clever method developed to overcome the problem that all the fundamental (000 - 001) absorption lines of CO₂ comprise a total effective bandwidth of only about 0.44 nm (centered around 4.256 μm -- see Chapter 2), a tiny IR fraction of the entire broadband solar spectrum [Christiansen, 78]. Thus raw solar energy, whose spectral distribution corresponds to a blackbody temperature of 5785 K, is an extremely inefficient direct incoherent pump for CO₂ lasers. The indirect pump instead heats an intermediate blackbody radiator surrounding the laser cavity to a temperature of about 1450 K, using either concentrated sunlight or perhaps nuclear energy [Christiansen & Insuik, 84]. The emission spectrum of that radiator then peaks at wavelengths around 4.3 μm , increasing the usable proportion by several orders of magnitude [Christiansen et al, 82]. High efficiency results because the intermediate blackbody constantly rethermalizes its emission spectrum, filling in the absorption profile left by lasing and thereby ultimately making all of the input energy available [Insuik & Christiansen, 84b].

The lasant, typically a mixture of CO₂, He and Ar, is optically thick for large lasers, because the gas container characteristic dimensions are then larger than the pumping photons' mean free path (average distance travelled before

absorption), which is of order 1 mm even for pressures around 1 torr [Yesil & Christiansen, 79]. Using pressures of order 10 torr increases gain [Insuik & Christiansen, 84a], because collisionally broadening the absorption lines increases the usable fraction of the blackbody emission spectrum. Gains on the order of 0.1 m⁻¹ have been demonstrated. Graphite is used for the intermediate blackbody; sapphire (Al₂O₃) is used for the laser tube itself, because it is about 85 % transparent to the pumping wavelengths, it is mechanically strong, and it holds up well under the 1450 K blackbody temperature (it absorbs heavily at wavelengths above 6 μm).

Heating is in fact the greatest problem for the indirectly pumped CO₂ laser. A lasant temperature of about 310 K maximizes gain; above 400 K, the lower laser level (100,020)_{I,II} becomes thermally overpopulated, degrading the inversion and eventually quenching lasing if hot enough [Golger & Klimovskii, 84]. All the various configurations proposed for large-scale space uses thus include provision for actively cooling their lasant. Typically the cavity takes the form of a counterflow heat exchanger, with jacketed or cored laser tubes; the lasant gas mixture, precooled by a purifying slurry of dry ice and acetone, flows past an opposing adjacent coolant (such as N₂), also cooled by dry ice or LN₂. Nitrogen is transparent to the pumping photons, so a variety of coolant/wall/lasant geometries are possible. The separate outputs of all these laser tubes are phase-locked by adaptive cavity length control, so they can be combined into one coherent, powerful beam by the transmitter optics [Lancashire, 82d].

Non-Planetary Laser Systems

Unlike the planetary lasers engineered in Part 2, high-power space lasers of the many types referenced above are already widely accepted as feasible even in the near future. Since an adequately detailed quantitative design for non-planetary laser transmitters capable of effective interstellar communication would consume a tradeoff study as elaborate as Part 2, here we can only posit their qualified feasibility based on the many high-power space laser studies. But such lasers are subject to their own peculiar engineering challenges; thus following the pattern established in the rest of this work, we now uncover some of the major problems a solar-pumped gas device might encounter. And to broaden further the range of potential system options, we suggest also some provocative approaches for their solution.

As explained earlier, CO₂ lasant mixtures must be kept from rising much above room temperature, a task made difficult if the laser tube is after all kept in a 1450 K oven. Remember that the reason for the intermediate blackbody was to increase the usable percentage of the broadband solar spectrum by shifting its emission peak to 4.3 μm . Although that works, clearly the spectrum is still inefficiently broad; all those other photons with wrong (non-pumping) wavelengths contribute to gas heating, either directly or by heating the container walls. One alternative way of improving the usable spectral fraction greatly even over this indirectly-pumped case would seem to be direct pumping which only allowed 4.3 μm photons to enter the laser system in the first place. The simple objection to this approach for CO₂ is that since the solar flux at 1 AU is not terribly strong and since the pumping wavelengths total only a minuscule fraction of that solar

spectrum, unreasonably enormous collector areas would be necessary.

But solar-pumped high-power lasers would most sensibly be stationed much closer to the sun than 1 AU, as suggested by Forward [84]. At the orbit of Mercury, for instance, the solar flux is 9 kW/m^2 , over 6 times stronger than at 1 AU. If about 10^{-6} of that is usable, and if the plug efficiency of the transmitter is 15 %, a 25 MW beam would indeed require a huge collector (radius of order 80 km if disk-shaped). The selective collector would consist of membrane panels surfaced by diffraction gratings (Chapter 7). With a blazed ruling unique for its position in the collector, each grating would reflect almost all of the desired wavelength band onto the laser apparatus, while scattering other wavelengths elsewhere. (A similarly selective, directional concentrator based on transmission, rather than reflection, holographic optical elements (HOEs) has been proposed for solar thermal rocket propulsion [Mickler, 85].) For longevity in the near-solar space environment, the thin-film membrane substrate would have to be a metal, or other crystalline or vitreous material.

Although the required collector area is vast, an interesting system interaction could turn that size to advantage. Because of the dispersive gratings, we consider the collector panels perfect diffuse reflectors normal to the sun line. If the overall transmitter system specific mass, dominated by its extensive sunlight collector, were kept to 1.26 g/m^2 , the radiation pressure of sunlight would balance the diaphanous spacecraft's gravitational attraction for the sun, so that it would not need to orbit the sun at all. Suspended photonically above its power source, such a device could be stationed anywhere on a $4\pi \text{ sr}$ sphere concentric with the sun, remaining in place and aimed at its receiver target continuously. Lateral, retargeting maneuvers and attitude

control about the sun-line axis could be done with ion engines. Both radial maneuvers and attitude trim about axes normal to the sun line could be achieved by controlling, quickly and with very low power, the reflectivity of individual sun-facing panels having liquid crystal coatings [NASA TB, 87]. Such non-orbiting transmitter stations would never interfere with each other astrodynamically, even if enough were distributed over the sun to devote one, or more, full-time to every colonized star system within 25 pc.

Although a photonically suspended spacecraft simplifies its own astrodynamics, pointing and tracking, the requisite 1.26 g/m^2 specific mass will not be easy to achieve. Forward [83] explains that sub- μm perforations in a photon sail can lightweight it without compromising its reflectivity. He also discusses in some detail thin-film manufacturing methods of achieving both this perforation and the thermal emissive properties necessary for near-solar use, and concludes that reducing the currently achievable 6 g/m^2 for solar sail film by an order of magnitude is worth investigating. Another obvious problem is that the laser station would clearly not be in free-fall. In fact, every part of it would feel a gravitational acceleration of 40 mN/kg , about 0.4 % of the acceleration at Earth's surface. This of course would affect its structural design, but a more fundamental issue is getting supplies to and from the spacecraft (and its construction site!). An arriving vessel would have to neutralize its solar orbital velocity along such a trajectory as to dock before the sun's gravity could pull it beyond recovery, neither a simple nor energetically inexpensive feat.

Even if combining photonic suspension with selective power collection for a direct-pumped CO_2 laser proves impractical, the advantages of non-orbiting stations might benefit other types of solar-powered lasers. All other schemes will in

general suffer predominantly from cooling penalties. Apart from cavity heating, the beam-combining and targeting mirrors (not to mention whatever modulator is used) will experience substantial thermal loads. Even absorbing only 0.1 % of a 25 MW beam results in a dose of 25 kW. Because the turbulence of actively-flowing heat transport fluids would compromise a mirror's fine-pointing accuracy (Chapter 7), we might expect the system mirrors to be cooled passively only (although running hotter would leave them more susceptible to laser damage from high beam power densities). Control optics such as resonator mirrors, couplers, modulators and targeters should in any case be structurally isolated by multiple-stage space-bearing actuators from the spacecraft disturbances caused by active cooling and attitude control systems.

A variety of compact radiator designs might accommodate the active heat rejection load of a high-power space laser. Chapter 9 referenced (and then avoided for the planetary laser, since they are vibration sources) several types. Developed for space nuclear power plants, most try to maximize total radiative area by using dust particles or liquid droplets. We favor here only long-lived, puncture-tolerant, contained systems. In the rotary disk radiator a liquid film spreads outward along the sloping inner surfaces of a rotating shell, to be collected for recirculation at its outer rim. An 8.5 m radius reference design for a 250 kW load achieves 0.58 kW/kg specific radiated power [Elliot, 85].

The RBMR, however, meets that same rejection load at 2.9 kW/kg, with a radius of only about 2 m. Indeed, because it is a two-phase system, in which a working vapor condenses on the inner bubble membrane surface before spreading centrifugally to be collected as a liquid, it can do the job of a single-phase system "with one-fourth the fluid mass and one-twentieth the mass flow requirements for the same operating

temperatures" [Webb & Antoniak, 86]. That of course means both mass and disturbance minimization benefits. A reference version scaled up to reject 17 MW still only has radius 7.5 m and total mass less than 2 MT [Coomes et al, 86]. Incidentally, fluid-filled rotary joints for thermal transport have been designed already [NASA TB, 85], so the RBMR should serve high-power space lasers well.

Finally we examine one crucial problem of contained-gas CO₂ lasers overlooked in detail by every published study we have referenced. As explained earlier, the typical choice of sapphire for laser tubes (with no hints about how to fabricate big ones) is made on the basis of its near-IR transmissivity, since the pumping photons enter the tubes laterally. But what seals off the ends of the laser tubes, through which the oscillating beam must pass on its way to the dynamically isolated cavity mirrors? Sapphire cannot work here since its transmissivity drops abruptly at 5 - 6 μm . The problem is one familiar to all CO₂ laser builders; most materials with high transmissivity at 10.6 μm are mechanically unreliable.

Materials demonstrating the lowest absorption at 10.6 μm are of several types [Marsh & Savage, 85] [OISPD, 84] [Savage, 85]: semiconductors (Ge); cubic crystals like halides (PbF₂, KCl); chalcogenides (ZnS, ZnSe, CdTe); the refractory oxides, nitrides, borides and carbides; the chalcogenide inorganic glasses, such as Ge-As-Se. Requiring antireflection coatings because of their high (> 2) index of refraction [Savage, 85], all of these are used commonly but none is suitable for really large, highly transmissive optics; chalcogenide crystals are conventionally ground up and hot-pressed into small blanks [Stierwalt, 75]. The best materials optically are the worst mechanically; soft KCl has an absorption coefficient at 10.6 μm of $1.4(10^{-4}) \text{ cm}^{-1}$. Polymers are in general poor in the FIR because of absorption features due to IR phonon

resonance with covalent bonds between their light elements (H, O, C, N) [Smith, 84]. In any case they are prone to laser damage [Dyumaev et al, 83] and would degrade in the space environment.

High total transmittance is really important for the laser tube end caps because they are intracavity optical elements (as noted earlier, the resonator mirrors would be outside the gain medium, analogous to the planetary laser). Losses due to their imperfect transparency add directly to those distributed losses which single-pass gain must overcome if laser oscillation is to occur. The more these elements absorb, the longer the gain path (and hence the tubes) must be to make up for it. It might indeed prove necessary to use cumbersome support methods for segmented caps of the available optically good, structurally awful materials reviewed above. Withstanding a few torr of gas pressure is not too challenging even for weak materials. But we suggest an alternative material, perhaps on the verge of realization for large optical elements.

As noted above, many cubic crystals are somewhat transparent to FIR wavelengths. The archetypal cubic crystal is diamond [Guy, 76]; in fact diamond has an extremely wide transmission spectrum [Wolfe & Zissis, 78], with absorption lines dependent on impurities [Woods, 84], and is "nearly transparent to FIR waves greater than 7 microns" [Brown, 87]. Diamond is receiving greater attention for electronic applications because of its extremely high electrical resistivity and thermal conductivity (its large-signal amplification figure of merit is 400 times greater than GaAs and 200 times greater even than InP). And of course its mechanical properties are legendary; albeit brittle, it is strong and stiff with an extremely low coefficient of thermal expansion (thermal shock factor 1000 times greater than sapphire), and natural lubricity comparable to Teflon. Most

promising, though, is the great recent progress in making useful diamond.

Synthetic (type Ib) diamond grit made with intense heat and pressure has been used for years as a machining abrasive. Then came diamond-like carbon (DLC) films, deposited most efficiently from gaseous CH_4 in an Ar atmosphere, typically using ion beams [NASA TB, 86]. DLCs are particularly useful because their composition (10 - 30 % H) and hence properties can be somewhat controlled, but they are not as clear in the IR as diamond. Finally, work by Soviet and Japanese scientists was confirmed in the US [Peterson, 86] that actual diamond films can be grown using chemical vapor deposition (CVD). Typically composed of separate microcrystals instead of a single matrix though, and limited so far in thickness to a few μm , such films are being vigorously researched. While it is not yet possible to buy sheets of diamond, it is reasonable to assume that continued work making these films (and similar films of cubic boron nitride (CBN), the second-hardest material known [Aerosp Am, 87]) can only yield improved IR windows, perhaps eventually even of unsubstrated diamond. Such components will greatly enhance the feasibility of large space-based CO_2 laser tubes.

Based on the considerable array of high-power space lasers already being planned, then, non-planetary lasers for interstellar communication should after all be feasible; but they will certainly not be easy to make. Because of their much smaller size, control systems to operate them should be attainable much more quickly than those needed for planetary lasers, and more conventional spacecraft subsystems can be employed for their housekeeping. Their limiting technology will instead be materials development. Of particular importance will be strong, stiff, lightweight structures, robust thin films, and large, heat-tolerant, transmissive

crystalline optics. A number of interesting system approaches warrant careful investigation to ferret out the fruitful ones.

The signal beam emanating from a high-power non-planetary laser is inherently more hazardous close to its source than one from a planetary laser; if distributed over a cross section of a few m^2 , for instance, a 50 MW beam is hundreds of times stronger than the solar flux even at Mercury. Thus such beams should only be wielded near the ecliptic with great care, as they could blind sensors and damage materials of other space systems. Nonetheless, acquiring such tools can avoid the eventual problem of having only a few interstellar data links available, by opening the possibility of dedicating full-time facilities to each active stellar target. Non-planetary lasers also broaden considerably the selection of potential laser sources beyond the planetary type explored in Part 2. Taking both kinds together and paraphrasing Freeman Dyson (Chapter 1), we conclude that there is no lack of laser systems available to any creatures which possess a desire to communicate efficiently around their galactic neighborhood.

References

Aerospace America Aerospace Materials "Slippery Coatings
Nearly as Hard as Diamonds" p 66 (Apr 1987).

K W Billman (ed)- Radiation Energy Conversion in Space
Progress in Astronautics & Aeronautics Vol 61
(AIAA, 1978).

Alan S Brown- "Diamonds Shine Brightly in Aerospace's Future"
Aerospace America Vol 25 No 11 (Nov 1987).

Walter H Christiansen- "A New Concept for Solar Pumped Lasers"
in [Billman, 78].

W H Christiansen & R J Insuik- "High Power Blackbody Pumped
CO₂ Lasers" in: Gas Flow and Chemical Lasers
ed by M Onorato (Plenum Press, 1984).

W H Christiansen, R J Insuik, R J De Young- "A Blackbody
Radiation-Pumped CO₂ Laser Experiment" NASA TM 84541
(Sep 1982).

E J Conway & R J De Young- "Solar-Pumped Laser Research"
presented at CLEO (24 May 1985).

E P Coomes, D Q King, J M Cuta, B J Webb- "PEGASUS: A
Multi-Megawatt Nuclear Electric Propulsion System"
Proceedings 21st IECEC (Aug 1986).

Russell J De Young, G H Walker, M D Williams, G L Schuster, E J
Conway- Preliminary Design and Cost of a 1-Megawatt
Solar-Pumped Iodide Laser Space-to-Space Transmission
Station NASA TM 4002 (Sep 1987).

K M Dyumaev, A A Manenkov, A P Maslyukov, G A Matyushin, V S Nechitallo, A M Prokhorov- "Transparent Polymers: A New Class of Optical Materials for Lasers" Soviet Journal of Quantum Electronics Vol 13 No 4 (Apr 1983).

David G Elliot- [op cit Chapter 9].

M J Fogg- "The Terraforming of Venus" Journal of the British Interplanetary Society Vol 40 pp 551-564 (1987).

Robert L Forward- Alternate Propulsion Energy Sources Air Force Rocket Propulsion Laboratory AFRPL TR-83-067 (Dec 1983).

- (84) [op cit Chapter 1].

A L Golger & I I Klimovskii- "Lasers Pumped by Solar Radiation (Review)" Kvantovaya Elektron Vol 11 p 233-357 (Feb 1984).

A G Guy- [op cit Chapter 8].

Paul F Holloway & L Bernard Garrett- "Utility of and Technology for a Space Central Power Station" in [Williams & Conway, 82].

Donald H Humes- "Preliminary Study on the Use of Lasers for the Transmission of Power" in [Williams & Conway, 82].

Robin J Insuik & Walter H Christiansen- (a) "A Radiatively Pumped CW CO₂ Laser" IEEE Journal of Quantum Electronics Vol QE-20 No 6 (June 1984).

- (b) "Blackbody-Pumped
CO₂ Laser Experiment" AIAA Journal Vol 22 No 9
(Sep 1984).

Richard B Lancashire- (a) "Laser Rocket System Analysis"
in [Williams & Conway, 82].

- (b) "Design Investigation of
Solar-Powered Lasers for Space Applications"
in [Williams & Conway, 82].

- (c) "Laser Power Conversion System
Analysis" in [Williams & Conway, 82].

- (d) "Laser Transmitters"
in [Williams & Conway, 82].

K J Marsh & J A Savage- "Infrared Optical Materials for
8 - 13 μ m --- Current Developments and Future Prospects"
Infrared Design SPIE Vol 513 Part 1 (1985).

Steven A Mickler- "Direct Use of Solar Energy for Orbital
Payload Transfer" Proceedings 22nd Space Congress
(23-26 Apr 1985).

NASA Tech Briefs- "Rotary Joint for Heat Transfer" p 155
(Winter 1985).

- "Depositing Diamondlike Carbon Films" p 139
(May/Jun 1986).

- "Liquid-Crystal Thermal-Control Panels" p 38
(Jan 1987).

Optical Industry & Systems Purchasing Directory

(Optical Publishing, 1984).

Ivars Peterson- "Diamond Electronics: Sparkling Potential"

Science News Vol 130 p 118 (23 Aug 1986).

Mark A Prelas, Frederick P Boody, Mark S Zediker- "An Aerosol Core Nuclear Reactor for Space-Based High Energy/Power Nuclear-Pumped Lasers" in [El-Genk & Hoover, 85] [op cit. Chapter 9].

J A Savage- "A Review of General Properties of Crystalline Materials for Infrared Optical Application" Recent Developments in Materials and Detectors for the Infrared SPIE Vol 588 (1985).

Alvin Smith- "An Investigation into Polymer Design and Synthesis for Infrared Energy Absorption" US Army Corps of Engineers Contruction Engineering Research Laboratory Technical Report M-345 (May 1984).

D L Stierwalt- "Low Temperature Transmittance of Materials for the Infrared" Long Wavelength Infrared SPIE Vol 67 (1975).

R Taussig, C Bruzzzone, L Nelson, D Quimby, W Christiansen- "Solar-Pumped Lasers for Space Power Transmission" AIAA Terrestrial Energy Systems Conference, Orlando (4-6 Jun 1979).

Willard R Weaver & Ja H Lee- "A Solar Simulator-Pumped Gas Laser for the Direct Conversion of Solar Energy" Proc. 16th Intersociety Energy Conversion Engineering Conference Atlanta Vol 1 p 84 (1981).

B J Webb & Z I Antoniak- [op cit Chapter 9].

M D Williams & E J Conway (eds)- Space Laser Power
Transmission System Studies NASA CP 2214 (1982).

M D Williams & L Zapata- Solar-Pumped Solid State Nd Lasers
NASA TM 87615 (Oct 1985).

William L Wolfe & George J Zissis (eds)- The Infrared
Handbook (Office of Naval Research, 1978).

G S Woods- "Infrared Absorption Studies of the Annealing of
Irradiated Diamonds" Philosophical Magazine B Vol 50
No 6 (1984).

Oktay Yesil & W H Christiansen- "Solar Pumped Continuous Wave
Carbon Dioxide Laser" in [Billman, 78].

- "Optically Pumped Carbon
Dioxide Laser Mixtures" Journal of Energy Vol 3 No 5
(Sep-Oct 1979).

CHAPTER 11

COST

Chapter Abstract - Projects like interstellar laser transmitters are so far beyond our current ability that familiar cost units would be meaningless applied to them. More productive is determining, through the prerequisite industrial infrastructure, the type of culture for which such projects would be acceptable. A mature interplanetary civilization, adept at space manufacturing and commanding large transportation energies, could construct and operate interstellar data links with only a small fraction of its productive economy. Unpredictable advances in such fields as material science, high-temperature superconductivity, and energy conversion would reduce the size, but not really the scale, of the required effort. Developing nanotechnological assembly, however, would drastically simplify and improve construction and performance.

Meaningful Cost

No feasibility study is complete without addressing the issue of cost. Many possible projects become in fact unfeasible (or at least not sensible), regardless of their technical merit, when the cost of implementing them is analyzed. Readers who persevered through Part 2 will know emphatically that the kind of interstellar lasers postulated therein would be fabulously expensive by current standards. Take for example just the active truss structures used throughout the planetary laser fleet. Considering how sophisticated the active members and active nodes must be, it is easy to imagine their development cost being measured conservatively in tens of \$M referred to today's manufacturing climate. Now although the fabrication cost of the subsequent members would be much less, there are still about 35(10⁶) such members in the fleet. And of course there is much more to the fleet than just a bunch of expensive truss members. Clearly Congress will not appropriate funds for such a project in the foreseeable future.

But we are far from ready to start construction anyway. Operating a planetary laser for interstellar communication requires control technology not yet available (Chapter 8), while operating a non-planetary laser for the same purpose requires materials not yet available either (Chapter 10). While both seem physically possible, both lie nonetheless beyond our grasp as yet, and will be practically realizable for us only after much progress. It is naturally tempting to conclude that the lasers' system costs would be prohibitively exorbitant even after, and partly because of, overcoming their remaining technical problems. Their sheer size, complexity and remote location, after all, seemingly embody entire economies' worth of engineering contracts. The fallacy of such a

premature conclusion resides in measuring necessarily future quantities with available, necessarily inappropriate standards.

To help illustrate how treacherous such chauvinism is when attempting to evaluate future cost, we adapt an example from Rood [85]. It happens that the non-organic (not combustion and not animals) mechanical energy usage in all of England in 1066 AD was carefully documented by the Dome's Day Book (an encyclopedic census) to consist of about 6000 waterwheels, each of which converted energy at about 2 hp. Now it would be tough to convince an 11th century Briton that a scant 920 yr later, thousands of people would fly every day, in several dozen machines, to and from his country. If astute in a pre-Newtonian sort of way, the Briton might point out that the required energy alone would make such a feat most unlikely. Explaining that devices called turbofan jet engines would indeed make it boringly common for a single 747 to consume 6 times as much power as produced by all of his country's power plants, would succeed only in convincing him of the ridiculousness of the whole idea.

Like the 11th century Briton, it is practically impossible for us to accept rationally the cost of building interstellar communication lasers. Certainly we have no meaningful unit for quantifying that cost. What, after all, might \$ 1 mean four centuries from now? Instead of imagining how much it would cost us to build an interstellar link, we can approach the problem in a more reasonable way by asking what kind of culture we would have to become before being able to do it. Our most sensible method of analyzing cost is thus to examine the infrastructure strictly required to make, emplace and operate interstellar lasers, realizing that that infrastructure should represent only a marginal fraction of the interplanetary culture undertaking the project. Over the past quarter century, NASA's budget has varied between 4.3 % and

0.75 % of the US national budget [Adams, 88]. If we assume that 1 % of our budget as a space-based human civilization were disposable for building a planetary CETI transmitter, we can imagine based on the sequel what the other 99 % of that culture would have to be like. We start to see what we will have become by then.

Infrastructure Cost

We discuss four broad cost categories underlying an ability to "do" interstellar lasers: development, material, fabrication and transportation. It would be a mistake to think that all the research and development precursor projects required by such devices must be amortized by them alone. Virtually every system outlined in Part 2 and Chapter 10 would have been developed for, tested and used in other kinds of projects long before the construction of interstellar lasers. Large actively segmented mirrors will be commonplace because of the enormous amount of astronomy a space-based civilization will do. Actively stiff structures and large AMCDs to orient them will be familiar already because of an inescapable need to control the attitude of all kinds of LSS. Embedded fiberoptic sensors and layered actuators will already be stock hardware. Spacecraft instrumentation and intercraft optical telemetry links will necessarily be highly-developed, too. Modular nuclear power plants of many types will have been refined for uses all over the solar system where concentrated, continuous power cannot be collected from the sun. Ion propulsion is even now well-characterized. And it is difficult to posit extensive solar system exploration and use at all, without adaptive artificial intelligence and optical computing.

Virtually the only aspects of interstellar lasers (as envisioned by this study) which would require "dedicated" development effort are those involving lasing per se. And as the literature cited in Chapter 10 shows, non-planetary high-power lasers will likely be developed for other purposes (power transmission and propulsion), again long before resources would be committed to an informational interstellar link. All that's left then is the Venusian laser itself, two aspects of which in particular require detailed study. The first is simple and can be decided easily within the next half-century: does the natural emission operate continuously and reliably enough, at high enough single-pass gain, to justify engineering a planetary resonator at all? Inevitably ongoing study of Venus will be able to answer this fundamental question in any case.

The second aspect involves what we may call the fine scale of planetary resonator operation. A small, experimental proof-of-concept resonator consisting of just two satellites would be used to verify that establishing a laser's coherence over thousand-km separations can be accomplished in fact (ironically though, the pointing loss tolerances for small reflectors are much more stringent than for the full-scale resonator stations -- see Appendix A7-10). That experimental equipment would then be used to study empirically the natural gain properties of the mesosphere, and to begin learning about transverse field variations of large-diameter planetary resonators.

After such preparation would come the bulk of detailed project development. Extensive human and machine expertise practiced in interplanetary construction of large space systems would bring together knowledge from related precursor projects to assemble and refine a real system design, optimizing it numerically throughout. Only then could the materiel, and

fabrication and transportation resources be arranged and committed.

Building a planetary laser would take large amounts of material which is not found naturally in Venus space, but must be brought from somewhere else. For instance, the fleet requires about 260,000 MT of beryllium for its mirrors, and about 50,000 MT of C/Mg composite for its active structures. Large amounts of fissionable uranium, refractory metals, C/C composite, lithium, and exotic semiconductors are needed for its power plants. Much copper is required for power trunks and EMT coils. Fittings and fasteners throughout the fleet are typically titanium, and silicon is needed for optical fibers and radiator surfaces. Xenon propellant must be resupplied periodically. Acquiring these diverse elements and the numerous others needed in smaller amounts is not trivial, since the solar system has not in general concentrated them together.

Virtually nothing would come from Earth by that time, because the interplanetary culture's industrial base could not have grown if imprisoned by a planetary gravity well. Not all the raw materials could come from the Moon either, an automatic but too glib presumption popularized by contemporary discussions of space industrialization. It is true that the Moon is rich in some elements important to industry, notably Si, O, Fe, Ca, Ti, Al, and Mg, and has substantial proportions of Na, K, Mn, P, Co and Cr as well [Adler, 86] [Taylor, 75]. These elements are not just lying around ready to be made into spacecraft parts, though, being instead locked up chemically in typically basaltic minerals. Nonetheless, the Si, O, Ti, Mg and Al our fleet requires might be of lunar origin, since those same elements will be important for all the precursor projects occurring in cis-lunar space long before the era of a Venusian laser. Certainly industries for extracting, refining

and launching these materials into space from the Moon will be economic fixtures by then, if not already defunct.

The Moon does have all the naturally occurring elements, but is extremely deficient compared to Earth in the lightest and most volatile ones (some of which are industrially critical), presumably because its small gravity could not retain them during thermal differentiation over geologic time. Typical relevant abundances are Li : 6.5 ppm, C : 7.2 ppm, N : 190 ppb, Be : 140 ppb, and B : 16 ppb. Abundances of the two lightest elements (H and He) are greater than expected (H : 1.6 ppm) because the solar wind has been impacting them into the lunar surface over the Moon's lifetime. Plans are being studied even now for strip-mining the minute quantities of ^3He from lunar regolith for use in fusion power plants [Phillips, 88]; preliminary estimates indicate that processing 4000 m³ of regolith per day to yield a few grams of ^3He would yield almost a tonne of H, and other trace elements as well. However, in general we may assume that enormous quantities of these elements (260,000 MT of Be, for example) will not be mined from the Moon.

Scenarios for extensive cis-lunar operations assume that lighter elements will be recovered from asteroids, the most common type of which (carbonaceous chondritic, CC) is enriched in precisely those elements lacking on the Moon. The total asteroid mass in the solar system is about $3(10^{18})$ MT, roughly 4 % of the Moon's mass [Hartmann, 83]. Although the inter-orbital transportation cost is higher to bring material from such lodes to industrial facilities in cis-lunar space, their surface launch cost is insignificant due to their minuscule gravity. Furthermore, some metallic asteroids represent highly refined metal lodes; thus a post-lunar inner solar system culture would use the transportation

infrastructure emplaced for CC recovery to supply many of its metal needs as well.

Most asteroids are of course located in the asteroid belt between the orbits of Mars and Jupiter. The belt is zoned compositionally, with stony asteroids dominating the inner edge, metallic ones comprising the central belt, CCs dominating the outer edge, and asteroids spectroscopically rich in organics inhabiting the extreme outer belt [Hartmann, 83]. For supplying a 1 AU industry, lower recovery costs would attend using the Earth-crossing Apollo asteroids instead, because their orbits have lower energy. The estimated Apollo population is of order 1000 having diameter larger than 1 km -- a substantial material source. Yet lower recovery costs would characterize Earth Trojans (if they exist), the asteroids presumed to occupy L4 and L5 in the sun - Earth gravitational system and named after those known to occupy Jupiter's stable libration points.

Mercury may turn out to be literally the gold mine of the solar system. According to the condensation/accretion theory of planetary formation confirmed so far by all available data, Mercury accreted from the portion of the solar nebula enriched by the refractory elements and compounds condensing first, including for example tungsten, and oxides of titanium and magnesium. Mercury is thought to be the most metal-rich of the terrestrial planets [Hartmann, 83]. With a surface escape velocity less than twice that of the Moon, and almost as airless, it could become an important source of rare metals for an interplanetary culture. Its extreme orbit location is not really even remote considering the large number of installations (whether for communication, propulsion or power) that such a culture can be expected to station there to take advantage of strong sunlight.

Material mined around the solar system should if practical be at least somewhat refined in situ. This cuts considerably the high energy cost of transporting the useful fraction to other orbits. While reducing this cost would probably lead a truly stellar culture to concentrate its industrial activity near the asteroid belt (Chapter 1), we can assume that planetary lasers will be affordable earlier in our history, when space industry still lingers around 1 AU. Unmanned, slow inter-orbital freighters could use argon ion propulsion, since rarer xenon would eventually be limited by a supply economy to special uses like station-keeping. (Atoms ejected at over 44 km/s along orbit tangents for transportation thrust are truly gone forever, since this exceeds even solar escape velocity for orbits farther out than Venus.) Ejecting process slag as reaction mass with EM launchers is another commonly proposed, albeit more polluting, propulsion method for large interorbital freighters. The Venus project can simply employ whatever emplaced inter-orbital transportation infrastructure already supplies the culture's larger needs. Source quarries would be the same, as would the destination terminals at 1 AU.

Raw materials combined and processed into factory stock would also be fabricated into the Venusian laser's component parts at already established facilities, for three reasons. First, the additional capital and operating costs to do so are just marginal (1 % in our scenario). Second, sending only finished parts to Venus minimizes the extra, dedicated transportation capacity required by that inter-orbital extension. Finally, the two important properties available cheaply to space factories are microgravity and hard vacuum. An inevitable consequence of using orbital vacuum in manufacturing processes is its eventual chemical degradation. Keeping this contamination and other debris out of Venus space altogether leaves its orbital environment pristine for the laser's operation.

A culture which could sensibly absorb the "exorbitant" cost of planetary or non-planetary interstellar communication lasers then, must be a truly interplanetary one which derives its energy from the sun and from nuclear plants, its material resources from the entire inner solar system, and its industrial strength from an economic web spanning at least Earth's double-planet system and the cities in its solar orbit. Capable of moving large payloads around in the solar system efficiently, such a civilization would undoubtedly be expanding its economy into the asteroid belt, and extending its reach into the icy outer solar system, where sources of both scientific knowledge and the volatile elements needed for biomass could keep it growing for millennia. Such an embryonic stellar civilization would unavoidably begin seriously to contemplate the prospects for extrasolar life, both finding it and causing it.

Shortcut Technologies

Although this study has tried to base its conclusions only on technology either already in hand or defensibly within reach, we cannot casually ignore the possibility of certain important breakthroughs which would make an interstellar communication project easier. The designs outlined herein would undoubtedly seem quaint to a culture practically capable of engineering lasers for such a purpose, although probably not as primitive in detail as our 11th century Briton's image of a 747. Because this chapter couples expense with historical timing, making the project easier means making it cheaper, which in turn means making it feasible sooner. In this section we thus examine the implications for planetary and non-

planetary communication lasers of potential advances whose timing remains unpredictable. We specifically omit the predictable progress, respectively, in artificial intelligence (Chapter 8) and material science (Chapter 10) already required in any case for developing such devices.

Non-essential material science advances would nonetheless enhance the laser project. Two specific improvements (mentioned in Chapters 8 and 9) that would directly and indirectly reduce spacecraft mass in the planetary laser fleet are high field-strength, high coercivity permanent magnets for use in EM actuators, and robust semiconductors with high Seebeck coefficients for use in TE power converters. Increasing component efficiency not only reduces their own mass in general, but also reduces the total power requirement and hence the power plant mass as well. Any improvements in the thermal properties and laser damage tolerance of materials for optics, non-optical members and fittings would enhance the endurance of non-planetary solar-powered laser craft, and decrease their specific mass (by permitting smaller sections) as well. It is reasonable, but not exactly predictable, that mature space manufacturing would achieve these and other incremental improvements.

In 1988, we have better reason than ever before to posit the grail of high-temperature superconductors as engineering materials. In a single year's time, certain perovskites have been made whose transition temperatures (at which they lose all resistance to electrical current) have been improved from a few K to well beyond nitrogen's boiling point (77 K) [SN, 87]; superconductive properties have been observed in some of the ceramics as high as 500 K [Peterson, 87]. Progress was also made achieving reasonable current densities through the new materials, although their granular, heterogeneous structure has proved variable and difficult to fashion usefully. It

would be hard to believe that space manufacturing methods will not benefit high-temperature superconducting materials, perhaps making them generally available to space industry. If such materials could be used in the power trunks and EM coils of the planetary laser fleet (particularly in space-bearing actuators and AMCD suspensor/ drive stations), its power requirements would plummet. Except for its ion engines, most of the fleet's power is consumed after all by resistive dissipation.

Similarly, just being the most appropriate nuclear energy source we have available so far does not guarantee that fission reactors will power planetary lasers. Potential future candidates for the kind of concentrated, endogenous, high-power energy conversion the fleet needs include muon-catalyzed "cold" nuclear fusion [Rafelski & Jones, 87] and antiproton annihilation [Forward, 83]. Both of these funded and promising techniques evolve heat as the intermediate energy form, which, then could produce electricity statically using the TE or THI conversion methods outlined in Chapter 9. The fusion method would require deuterium and tritium as fuel, while the mirror-matter method would require long-lived antihydrogen ice, produced by an accelerator factory. In either case the necessary production infrastructure would already be supplying these exotic fuels to power plants operating throughout the settled solar system.

Thus high-temperature superconductors would reduce the power plant capacity required to run the planetary laser fleet, and advanced energy sources would reduce the mass and perhaps the size of those plants as well. Improvements in engineering materials would reduce both. Consequently any of these breakthroughs might reduce the laser system project's cost, making it feasible somewhat sooner. But none would change very much the scale of effort necessary to undertake the project, the experience required to pull it off, or its

operational performance. All that could accomplish such qualitative improvements over the scenario we have outlined so far is a fundamental revolution in manufacturing technology, which might indeed occur before the era of interplanetary civilization.

The history of technology is partly one of learning to effect more and more subtle changes on smaller and smaller assemblages of material. The end result of this progression, perhaps by now only generations away for us, must be an ultimate chemical ability to manipulate matter an atom at a time. Properly termed nanotechnology because a typical molecular dimension is of order nm, this ability will irreversibly increase both the prospects for, and threats to, the continued evolution of Earth life. Treating the myriad subtleties underlying the concept of nanotechnology, as well as the strong reasons to suspect its imminence, is way beyond the scope of this study; we thus refer readers for all such background to the basic text in the field [Drexler, 86], which contains much seminal analysis and further technical references.

The reality of nanotechnology has been demonstrated by every living thing over several billions of years. Proteins (the basic structure and machinery of life), in the form of enzymes and (in complex cells) ribosomes, process energy and assemble more proteins by manipulating matter reliably, an atom or molecule at a time. Proteins perform according to instructions implicit in their own molecular structure, encoded for ultimately by the genetic fine structure of self-replicating nucleic acid molecules. Life uses all this molecular machinery to reproduce microscopic unit cells quickly (exponentially) into vast numbers, in complex cases making up tissues, organs, organisms and therefore societies.

The premise of designed nanotechnology includes developing assemblers, machines only a fraction of a cell's size, which can put together and take apart molecules according to instructions. Those instructions are transmitted and processed by nanocomputers, molecular devices whose essential logic and structure might even be based on mechanical linkages. Because they are so small, vast quantities of both assemblers and nanocomputers (produced in short times by replicators) work together to achieve macroscopic effects. The engineering utility of such nanomachinery derives from two important distinctions between it and "natural" life. First, not constrained (as all protein-based life forms have always been) to one type of chemistry, they can be more compact, more efficient, more robust, and build arbitrary structures from arbitrary materials according to the laws of atomic bonding. Second, nanocomputers can contain, in volumes orders of magnitude smaller than a single cell, amounts of information and processing available biologically only to organ-sized collections of separate cells.

The challenging and enormous task of achieving nanotechnology depends on two fundamental abilities, both currently funded largely by defense interests: nanomanipulation and artificial intelligence (AI). Only with machines to perform detailed design and analysis can we hope to make other machines which can study, program and build to atomic specifications. Following a bootstrapped design process, nanotechnology will be able, shortly after the advent of the first assemblers, to develop nanomachines capable of building anything allowed by the constraints of physics, to atomic specifications, in any amount until their program stops them or they run out of raw material. Being molecular von Neumann machines (self-replicating devices which can tackle any job by producing sufficient copies of themselves), assemblers will indeed revolutionize manufacturing.

To give some idea of the range of abilities made possible by nanotechnology, we list some of Drexler's examples: flawless fibers of carbyne or diamond, to enable formerly impossible structures; atomically precise "machining" of AI-optimized hardware designs, to allow unprecedented lightweighting and material economy; desirable foods manufactured according to natural molecular patterns directly from C, O, H, N, P and trace elements; hardy disassemblers that could tunnel deep below the Earth's surface or in other inaccessible places, recording what they found; insidious nanoweapons that could selectively, quickly and incontrovertibly destroy anything they were programmed to, from a material threat, to a race of people, to the ecosystem of a planet; "active shields" to prevent such devastation; life extension and health maintenance through cell repair machines restoring damaged macromolecules, detoxifying poisons, and disassembling pathogens.

Clearly the advent of programmed nanotechnology will determine, endanger, and liberate the future of human existence as deeply as nuclear power, electricity, and even fire have done in the past. For a penetrating review of some possibilities, refer again to Drexler. Although the energetics of spaceflight will obviously not be affected, virtually all its hardware will benefit. Certainly optimal lightweighting, and homogeneous and exotic materials produced quickly and cheaply in enormous quantities, will enhance all space operations. Devices like adaptive environmental suits, and cell repair machines to undo radiation damage, will promote directly the human settlement of space. The specific advantages of nanotechnology for building and operating interstellar lasers are legion.

Assemblers will grow structural components lightweighted beyond present belief by omitting material everywhere that

numerical models indicate it is superfluous, resulting in striking mass savings and consequent performance enhancements throughout either a planetary or non-planetary fleet. Thus, for instance, solar photonic suspension using perforated HOEs will become realistic. The complex innards of active structural members will be assembled inside them, like bottled model ships. Electrical connections and composite materials will be bonded as reliably as any intrinsically homogeneous substance. Furthermore, all fabrication tolerances will perforce be atomic. Microroughness of optical segments, for example, will be due to the dimensions of their surface atoms alone, since those atoms' placement will be individually deliberate. The millions of sensor and telemetry instruments will be identical, consistent and truly accurate. Large sapphire gas tubes for non-planetary lasers will be made monolithically, with integral, optically perfect diamond ends. And with nanoassembly, of course, these and all other components can be made precisely as easily and cheaply as they can be made at all.

But by far the most exciting benefits will be operational. Using nanotechnology, each spacecraft will be less an assemblage of parts than a single machine, monitored at the atomic level by nanofixers. Under interfaced direction of the craft controller, these devices will effect any repairs imagineable in situ, supplied with materials when necessary by logistics butler spacecraft. Nanorepairs will proceed continually, precluding noticeable material degradation. And macrorepairs (as for instance after meteoroid impacts) will begin immediately after an accident, restoring damaged parts quickly and obviating gross replacement. Nanocare will make the spacecraft virtually indestructible. Finally, the controlled ability to grow new parts (and recycle the atoms from reabsorbed ones) has direct value for an adaptive fleet intelligence. Rather than carving its activity patterns

out of an expensively redundant optical processing network, the controller will be able actually to develop the nerve pathways it needs, constantly refining the material armature of its consciousness.

Unlike breakthroughs in conventional engineering materials, superconductivity or nuclear power, nanotechnology will totally redefine the scale and the cost of efforts to design, construct and operate interstellar communication lasers, simultaneously improving greatly their achievable performance. A nanotechnological civilization taking on such a project would then be one already materially rich beyond any avarice imagined in history, commanding almost any arrangements of matter to study, mimic and alter the natural world, channelling and controlling its energies. That civilization, one that we may quite possibly become, would be able and anxious to expand out into the galaxy, remaking itself in startling ways to which we now turn.

References

Laurence J Adams- "Civil Space Program in Crisis" Aerospace America Vol 26 No 2 (Feb 1988).

Isidore Adler- The Analysis of Extraterrestrial Materials (Wiley & Sons, 1986).

K Eric Drexler- Engines of Creation (Anchor, 1986).

Robert L Forward- (83) [op cit Chapter 10].

William K Hartmann- [op cit Chapter 2].

Ivars Peterson- "Signs of a New High in Ceramic Superconductivity" Science News Vol 132 p 356 (5 Dec 1987).

Paul Phillips- (Eagle Technical Services, Houston) personal communication (Feb 1988).

Johann Rafelski & Steven E Jones- "Cold Nuclear Fusion" Scientific American (Jul 1987).

Robert T Rood- "The Nature of Technological Civilizations" The Search for Extraterrestrial Intelligence (ed by K I Kellermann & G A Seielstad) Proceedings of a Workshop at National Radio Astronomy Obervatory, Green Bank (20-22 May 1985).

Science News- "Science News of the Year" references articles throughout the year Vol 132 p 402,405 (19 & 26 Dec 1987).

Stuart Ross Taylor- Lunar Science: A Post-Apollo View --
Scientific Results and Insights from the Lunar Samples
(Pergamon, 1975).

CHAPTER 12

INTERSTELLAR TRANSPORTATION

Chapter Abstract - Augmented planetary and non-planetary lasers could transmit as much data reliably to 25 pc as their optical carriers could ever carry, permitting up to 10^{20} b/yr to be sent to other stars. Nanotechnology will enable extremely small factories, propelled photonically across the void in only years, to build large, matched, far-field receivers from indigenous materials. The molecular cloning enabled by biostasis, nanorecording and cell assembly can then occur over interstellar distances using the laser link. If humans survive the transition to nanotechnology, redundant techniques for secure transmission of genetic and cognitive information from star to star make virtually certain the human metamorphosis into a galactic culture.

The previous 11 chapters have set a stage of practical interstellar communication at high data transfer rates. To the extent possible with such limited analysis and design, we have established several key conclusions. Apparatus to send a lot of information with highly-modulated electromagnetic carrier signals among hundreds of neighbor stars is not simple, nor is it exactly possible or affordable for a civilization at our stage of technological development. Having exposed the toughest basic technical problems, we nonetheless see several promising directions for progress which, certainly within centuries, would enable us to realize both non-planetary and planetary lasers to perform the mission. For the kind of interplanetary industrial culture we might almost momentarily become, therefore, the ability to communicate efficiently across interstellar space seems feasible.

In proposing such a completely new human ability, we incur an ethical duty to explore the kind of world it might usher in. Too often, technical prowess exercised either for expedient superiority, or for its own intrinsic thrill, neglects its shocking and sometimes irretrievable ramifications. The classic example from our own century is nuclear weaponry. Developing and using this powerful technology were such distinct activities that, only years after two cities had been obliterated, could Robert Oppenheimer admit, "In some sort of crude sense which no vulgarity, no humor, no overstatement can quite extinguish, the physicists have known sin; and this is a knowledge which they cannot lose" [Pell, 88]. It is a knowledge currently approached by biochemists learning to manipulate the chemical codes of life, and a knowledge awaiting venomously the programmers and molecular designers of the future nanotechnological era [Drexler, 86]. But avoiding such progress is not a viable course; as Drexler points out,

suppressing new technologies is at best ineffectual in a world anything less than perfectly totalitarian.

Only the foresight encouraged by open discussion of the potential changes new technologies introduce can mitigate the technical and social risks they bring. Thus no modern feasibility study can be responsibly complete without addressing the meaning its conclusions might hold for the known world. A particular synergy of topics already discussed in this work will, if after all possible, transform our world and perception profoundly, with a certainty assured by lower bounds on evolution. In one of the most chilling comments on the inevitable changes wrought by an available nanotechnology, Drexler says, "If the eons-old evolutionary race does not somehow screech to a halt, then competitive pressures will mold our technological future to the contours of the limits of the possible." In this the last chapter of our interstellar communication fugue, we assume finally that everything discussed earlier is possible; in thus pulling out all the stops, we dutifully establish limits and probe their contours to gain foresight. Combining the strains of nanotechnology and efficient interstellar data transfer leads directly to a possible future more exhilarating than any hinted by the theme which began Chapter 1.

Augmented Transfer Rate

Chapter 4 concluded, based on a reference Venusian laser design and reasonable link assumptions, that up to of order 100 Mb/s could be reliably transmitted among several hundred neighbor stars. Chapter 10 showed how other, non-planetary laser systems could match this performance. But with

relatively modest augmentation, the reference system could establish a link capable of data transfer at rates as high as the theoretical modulation limit of an optical carrier. Figure 12-1 shows the result of one set of assumptions (by no means the only set which can achieve that result).

As shown by the link equation (A3-3.7), increasing the laser cavity diameter and increasing the receiver diameter have the same effect and are therefore algebraically interchangeable. We keep the same far-field 1 km receiver diameter, but now take a laser cavity diameter of 5 km, using to maximum advantage the available planetary source (Chapter 7). We further assume the coupler and modulator hardware can effect 90 % beam throughput (instead of the 21 % accepted earlier to simplify pointing control). We still assume a human-built, bandwidth-matched receiver with 3.3 Hz channels, but now having an enhanced receiver degradation factor of only $\Delta R = 3.5$, closer than 5 to the theoretical limit of 3.1 (Chapter 3).

Even with such modest alterations, the link capacity can easily exceed any ability to modulate the beam. That is, even presuming wideband optical data processing and optically-mediated modulation methods not yet feasible for a large and powerful IR beam, there would be no point in modulating it at a frequency greater than about 10 % of its carrier frequency (a rule of thumb from optical communication theory [Glass, 87]). For 10.6 μm radiation, that limit is near 2 THz, indicated in Figure 12-1. So with only modest augmentation, targeting a spot the size of Mercury's or Venus' orbit is unnecessary; targeting a spot even as large as Saturn's orbit can permit the same transfer rates discussed in Chapter 4.

This result means two things. First, an optical interstellar link can accommodate modulation rates as high as

ever will be feasible for the carrier frequency. Second and just as vital, "excessive" augmentation lets these high transfer rates be achieved with K beyond 20, and hence at a much lower bit error probability PE than 10^{-9} . So given a bandwidth-matched link, accurately targeted laser beams (produced either by planetary resonators or high-power space-based sources) can transmit data, with almost arbitrarily high reliability, among neighboring star systems at rates as high as any optical modulation technology will ever allow. With a continuous duty cycle, that means that such communication links can transmit up to hundreds of exa-bits per year (order $100(10^{18})$ b/yr) among those star systems. Now if we had such an ability, what would we really use it for?

Biostasis and Molecular Recording

Nanotechnology promises to be an exceedingly dangerous tool, capable if used offensively or accidentally of consuming life on Earth with greater finality and quickness even than a nuclear holocaust. It also promises formerly incredible benefits of material wealth and health, and material engineering limited only by intelligence and the laws of chemistry. Long before these extremes, however, nanotechnology promises to be denied, feared, misunderstood, and misrepresented. In precisely these uninformed reactions resides its gravest danger, because when the first assemblers are made, nanotechnology will burst upon the world more abruptly and importantly than any other technological tool in history, whether or not the world is adequately prepared. The entire sequel is based on principles of nanotechnology which we cannot explain or defend in detail here. In developing the present thesis, we will however repeatedly focus

in on increasingly more specific implications of molecular engineering. For essential analysis, critical readers must consult Drexler [86].

Molecular computers will occupy about 1/500 the volume of a typical mammalian cell, yet contain vastly more information than the cell's own DNA --- enough to recognize and record the type and position of all the macromolecules found in normal cells --- and control a hierarchy of perhaps thousands of much tinier repair machines. Powered by the same type of chemical energy that runs the cell itself, each will "perform a thousand computational steps in the time that a typical enzyme takes to change a single molecular bond", deciding on molecular changes to execute within the cell. Together they will restore it to proper balanced function, assemble replacement parts from nutrient stores, repair damaged genetic material, remove toxic substances, sense, encode and transmit to a larger computer the biochemical fine-structure of the cell, or effect biostasis by blocking metabolic activity and erecting a preservative molecular scaffold among macromolecules.

While such operations will incidentally enable unprecedented and precise control over organic disease and degeneration, we concentrate here on the implications of biostasis, molecular recording and assembly. Chapter 8 already reviewed some basics of neurophysiology. Every shred of evidence amassed by science so far indicates that personality, memory, individual response and thought patterns are all encoded materially in the continually evolving but finite anatomical and molecular fine structure of the nervous system. There are about 10^{15} synapses in the human cortex [Changeux, 85] [Churchland, 86], and at least 45 neurologically active hormones have been isolated in the human brain [Bergland, 86]. By all indications, our unique structural patterns of neuronal

interconnection, synapse location, and chemical concentration are what make each of us uniquely complex.

The critical difference between expiration (clinical death) and dissolution (irreversible brain death) is that unarrested metabolic changes after the former eventually disrupt the brain's fine structure to cause the latter. Nanotechnological biostasis can intercede to preserve the fine structure, forestalling dissolution until physiological and biochemical repairs can reverse expiration [Drexler, 86]. Dissolving stasis by restoring proper fluid and electrolyte balance would revive the patient, just as flushing anesthetic from brain tissues allows interrupted consciousness to recommence. Again, while such biostasis will incidentally reduce the risk of sudden death from traumatic injury, allow long space journeys in "suspended animation", and even permit a kind of one-way "time travel" (delaying into the future the rest of one's life), we concentrate here on biostasis' ability to allow structural molecular inventory.

As they effect biostasis, intra- and intercellular nanomachines can transmit the types and relative locations of all macromolecules, as well as fine-scale anatomy, to waiting storage computers. If using the sheathed carbyne-rod system of mechanical nanocomputers, such signalling can occur through the patient's own blood vessels. Although the enormous amount of cellular standardization would reduce drastically the sheer amount of information such an inventory would need to record, an aorta-sized bundle of signalling fibers "can in less than a week transmit a complete molecular description of all a patient's cells to a set of external computers" [Drexler, 86]. That external record now encodes the entire current state, both the physical manifestation and consequent mental being, of the person's life. Although cognitive encoding might be used to build a tiny nanocomputer capable of modeling accurately the

person's brain for a variety of uses, we concentrate here instead on ametabolic cellular assembly.

Now if nanotechnological assemblers can arrange atoms in any desirable way possible, and if under nanocomputer control they can fabricate replacement parts for functioning cells, then they can also assemble functioning cells entirely from raw stock, following directions just as the cell's own machinery would. Specifically, they can produce cells and arrange them ametabolically (quiescently) into tissue, organs and finally a complete organism according to the exact molecular instructions recorded during stasis. Then when stasis is dissolved, what will awaken will be an exact molecular duplicate of the inventoried person. This startling ability to clone people, with their minds intact, from water and a few pounds of chemicals, is an inevitable, albeit undiscussed, result of Drexler's analyses of cell repair, biostasis and nanoassembly of extinct life forms.

Clearly societal concepts of individuality, opportunity and death will be shaken to their foundations by molecular cloning. While it is difficult for us to imagine carefully either the horror or the glory made possible by such a cultural transformation, it is promise, and not fear, which lures us on. People suffering irrecoverable death will be "reawakened" (less their final memories, of course) using their most recently recorded personal data, changing the human perception of risk. People will install their minds in new and different bodies, designed the way they want them. People with an identical starting set of memories, skills, and values will lead simultaneous or sequential lives, perhaps separated by enormous distances in time and space -- or perhaps not. Updating that most precious of all possible personal information, people will attain essentially a cognitive immortality, at least until their minds reach a limiting

capacity that no one yet can define. Finally though, we concentrate on only one of the possible uses for copying unique human structural information. Referring back to Chapter 3, we recall that to a communication system, all information is just data, even if the link is interstellar and the data encode a human life.

The Transporter

Nanoassembly makes possible extremely light, efficient photon sails (Chapter 11). Laser pressure can accelerate such a sail to a relativistic coasting speed able to reach another star system in only a few years. Drogue sail segments would separate en route, decelerating the sail-rigged payload with reflected laser light [Forward, 84]; or onboard disassemblers could consume the sail, fabricating with its atoms an accelerator to neutralize the payload's interstellar speed [Drexler, 86]. The payload would not be an enormously heavy, peopled starship, but rather a tiny microcraft just a few μm^3 in volume, stuffed with molecular instructions for many types of assembly and operation. This nanoseed would target a source of raw materials in the alien star system, assembling therewith a fleet of reconnoitering spacecraft and a station to receive "further instructions from home, including plans for complex devices."

If that receiver is a bandwidth-matched terminal for the type of laser link we have proposed with this work, then the "plans" could be those for the most complex "devices" we know, ourselves. Indeed, given a modulation-rate-limited interstellar data transfer channel and molecular recording and assembly biotechnology, using both together for the deep-space

transportation of human beings would seem inevitable. Assemblers could reconstitute a human clone as easily in another star system as next door at home, given the same information.

Before the ultimately comparative molecular anatomy which nanotechnology will allow, no one can specify with any authority even the order of magnitude of digital bits required to encode a human life, so we cannot yet determine how long an Eb/yr augmented laser link would take to send an encoded person. However, we can see why the problem is solvable. Although the number and location of synapses and chemicals which make us unique represent an awesome and potentially daunting amount of data, clearly neither the synapses nor the chemicals can be distributed truly randomly, so there must be severe constraints on their possible permutations.

Over 105 human genes [Weiss, 87] make up our 46 chromosomes, the smallest of which contains a sequence of order $5(10^7)$ quaternary bits (DNA base pairs) [Hood, 87]. And yet it is well-known that human DNA is over 98 % identical to chimpanzee DNA. Thus the genetic code is necessarily almost identical from human to human. The fine anatomical and biochemical structure of human brains must also conform to general "exclusion" rules for them to be human. Once those rules are elucidated, they can be sent along with the original nanoseed to a target stellar system; subsequently, only the (much more manageable) difference data between that standard recipe and a given individual need be sent with the laser link in order for the distant assemblers to reconstitute that person intact. A complete difference-data set would include unique genetic, neuronal, and somatic biochemical information (because the history of a given individual's recombinant antibody diversity, for example, is as unique as his mental memories).

Serious discussion of the quantitative difficulties and societal implications of cloning people is not original [Clarke, 84], nor even is the concept of using lasers to transmit life-descriptive information across interstellar space (one estimate supposes that 0.1 Eb could encode an individual) [Rather, 87]. Feoktistov [87] recognized the inherent advantage of deep-space electromagnetic data transfer over physical transport. That is, one of the "less" resolved problems of relativistic spaceflight is that interstellar space is not empty. Colliding with dust and gas, not to mention a real Oort particle, while moving at relativistic velocity could result in catastrophic vaporization, unacceptable erosion, or at least excessive induced ionizing radiation. Proponents of physical interstellar travel have suggested predictably baroque feats to preclude direct collision, but using electromagnetic radiation as the carrier sidesteps this issue entirely.

What is original in this work is the idea of combining feasible technologies for interstellar lasers with molecular assembly, providing the means both to send descriptive data among star systems and to reconstitute it into a human being once received. Together these technologies comprise the closest thing to a "transporter" that we are likely ever to have.

Manipulating, across space, information extensive enough to allow assembling an intact human individual is understandably unsettling, fundamentally much more frightening than sending movies and equations to aliens. The point-to-point nature of the laser link, which allows such great data transfer in the first place, also protects it largely from "unauthorized" interception, though. No receiver outside the targeted spot could tap a meaningful signal. Also, sending data is not the same thing as sending people themselves.

Should something go wrong with the link at any time, error detection intelligence would prevent assemblers from attempting to reconstitute a faulty or incomplete data set. And the "original" version will still exist at home in any case. This transporter beam, albeit much slower than the ones in science fiction and justified only for enormous distances, will be much safer.

We might well ask, if entire libraries can be recorded in tiny nanocomputers [Drexler, 86], why we would bother to transmit personally descriptive information via lasers. After all, if receiver nanoseeds themselves can arrive intact after relativistic trips, having successfully avoided collisional destruction, then why not just include in them the datasets representing entire people? At relativistic velocities, the trip time would take not much longer than a laser signal itself, and all the information would arrive simultaneously. Indeed this method of interstellar transportation also seems workable, although without a laser link no further datasets could be transmitted to the same target until another microcraft was launched. Throughout this work, we have tried where possible to bolster conclusions of feasibility by demonstrating alternative ways of accomplishing desirable ends. That theme recurs here, too. Given molecular recording and assembly, at least physically encoded interstellar transportation seems inevitable. And given in addition efficient lasers to link stars informationally, datonically encoded transportation seems inevitable also. Furthermore, even if laser-pushed relativistic propulsion proves unusable, more conventional nuclear-powered electric propulsion could still deliver payloads to other star systems, although such trips would take a few hundred years [Aston, 87].

Effective, near-lightspeed interstellar transportation for humans seems as robust a concept as nanotechnology itself,

limning some contours of the possible. As stated before, this study is not the proper place to argue for the inevitability of nanotechnology; but upon reviewing available analyses in the nascent field, the fundamental issue that emerges is not whether it will become available, but rather whether we can keep from destroying our ecosphere with it before we have controlled and adapted to it, when it becomes available. If we survive, literally a universe of opportunity will open to our species.

Freeman Dyson has for quite some time maintained that moving off Earth and into space, particularly into the Oort cloud defining the fringe of our solar system, will cause speciation, producing eventually different descendant species from our human stock [Finney & Jones, 83]. Nanotechnologically studying and designing ourselves will in fact bring speciation under willful control; humans will be able to monitor, suppress or cause evolutionary adaptations. More vital to our future as a galactic progenitor civilization than mere genetic chemistry, though, is the evolution in social intelligence which nanotechnology and feasible (but always channel-choked) interstellar travel will force. It is this abrupt, quantized evolutionary transition in human history and the history of life in the galaxy (virtually a textbook case of Gould's "punctuated equilibrium" model of species development) that we cannot really see beyond until we experience it in fact. But now is not too soon to start trying.

In an interstellar culture, who chooses which resources get consumed and which left wild? Who controls the precious information links among the stars? Who chooses which individuals of the trillions living in some solar system have their datasets transmitted across the void to other islands in space? For those few, life and growth will have entirely different meaning even than for the multitudes remaking

themselves within one stellar system. While all human cloning will be something vaguely like having children that could benefit from every detail of a parent's experience, interstellar replication will be yet grander. Some humans will thereby metamorphose into a new and multiple life form, spreading their consciousness outward into the galaxy as they incidentally seed it with human progeny. They will not do it for escape, nor for material gain, nor even for scientific curiosity, but simply because they can. They themselves will spawn, and be, the galaxy's extraterrestrial intelligence. N must become huge.

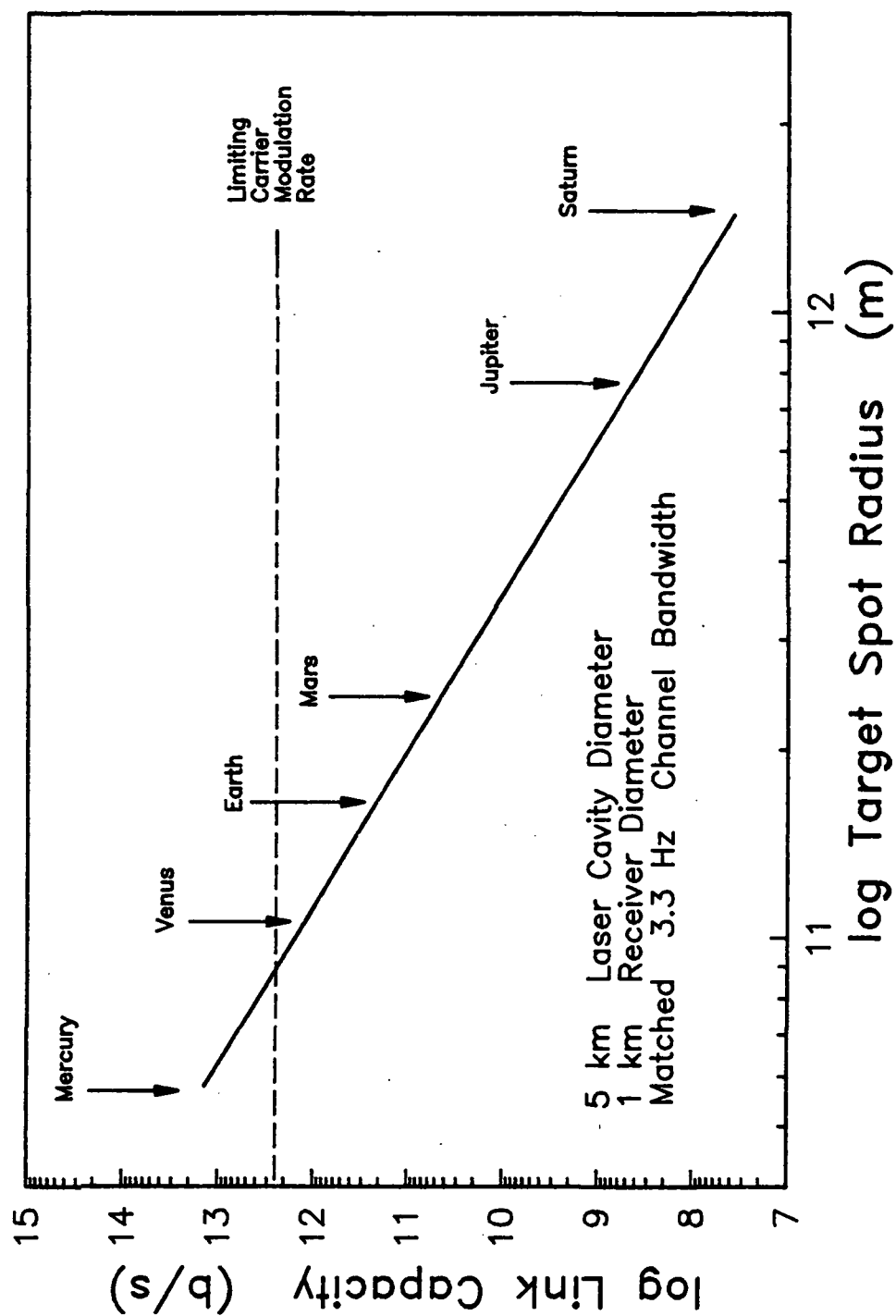


Figure 12-1 Limiting link performance, showing reference target orbit sizes.

References

Graeme Aston- (87a) [op cit Chapter 9].

Richard Bergland- The Fabric of Mind: A Radical New Understanding of the Brain and How it Works (Viking, 1986).

Jean-Pierre Changeux- [op cit Chapter 8].

Patricia Smith Churchland- [op cit Chapter 8].

Arthur C Clarke- Profiles of the Future (Warner, 1984).

K Eric Drexler- [op cit Chapter 11].

Konstantin Feoktistov- "A Flight to Stars" presented at 8th SSI/AIAA/Princeton Conference on Space Manufacturing, Princeton (May 1987).

Ben R Finney & Eric M Jones- "From Africa to the Stars: The Evolution of the Exploring Animal" 6th SSI/Princeton Conference on Space Manufacturing (May 1983).

Robert L Forward- (84) [op cit Chapter 1].

A M Glass- "Optical Materials" Science Vol 235 p 1003 (27 Feb 1987).

Leroy Hood- (California Institute of Technology)
interviewed in Omni Vol 10 No 2 (Nov 1987).

Claiborne Pell- "First Word" Omni Vol 10 No 5 (Feb 1988).

John D G Rather- (Kaman Aerospace Corp) quoted in "Laser
Clones" Omni Vol 9 No 4 (Jan 1987).

Rick Weiss- "First Human Genome Map Completed" Science News
Vol 132 p 245 (17 Oct 1987).

EPILOGUE

The thesis defense seminar of this work took place on 31 March 1988 at the University of Maryland. Professor Jean-Paul Richard in particular raised issues that should provide the starting point for any further study of this topic.

Equation A3-3.7, by folding in the heterodyne signal-to-noise power ratio, assumes sufficient photons in the signal stream to preclude single-photon statistics as the dominant noise contribution. In fact, the total number of received photons (for the reference design, of order 10^8 ph/s for a 1 km diameter receiver stationed within a central Airy spot equal in size to Mercury's orbit) is small enough to assure that the signal will be shot-noise-limited. That is, the minimum noise must equal the square root of the number of photons. Requiring $K = 20$ then limits data transfer to less than 1 kb/s. It appears that even the augmented planetary laser (using a full 5 km diameter cavity, and the other assumptions of Chapter 12) would be limited to rates less than 100 kb/s.

Such data transfer rates would of course be useful for some forms of CETI, but communicating with human stellar colonies, particularly in the manner explored by Chapter 12, depends on much higher data rates. Three approaches to achieving those higher rates need to be investigated.

First, alternative encoding schemes (such as bit repetition) may exist for keeping PE below 10^{-9} without requiring $K > 20$. Second, knowing the receiver's stellar orbital parameters might enable target spots much smaller even than Mercury's orbit. Receiver diameters larger than 1 km are probably feasible also. The number of received photons could then be increased substantially, with system performance limited by mechanical (pointing and mirror) abilities.

Finally, non-planetary lasers, as outlined in Chapter 10, could produce beams containing several orders of magnitude more photons than those of planetary lasers. The shot-noise limitation thus provides a further incentive to study such high-power laser systems as a viable means of establishing efficient interstellar communication links.

ENGINEERING PLANETARY LASERS FOR INTERSTELLAR COMMUNICATION

by

Brent Sherwood

Thesis submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Master of Science
1988

Examining Committee

Professor Bruce K Donaldson
Professor Michael A Coplan
Professor Bruce S Berger
Professor Jean-Paul Richard

NASA Technical Advisor

Dr Michael J Mumma

© Brent Sherwood 1988

All rights reserved.

Report Documentation Page

1. Report No. NASA CR- 180780		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Engineering Planetary Lasers for Interstellar Communication				5. Report Date May 1988	
				6. Performing Organization Code	
7. Author(s) Brent Sherwood				8. Performing Organization Report No. UMAERO - 88 - 5	
				10. Work Unit No.	
9. Performing Organization Name and Address University of Maryland College Park MD 20742 Department of Aerospace Engineering NASA Goddard SFC Code 693				11. Contract or Grant No. NGT 21-002-823	
				13. Type of Report and Period Covered Contractor Report	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington DC 20546-0001 Code 693, NASA Goddard Space Flight Center Greenbelt MD 20771				14. Sponsoring Agency Code	
15. Supplementary Notes Thesis submitted to the Faculty of the University of Maryland in partial fulfillment of the requirements for the degree of Master of Science, 1988. Work sponsored under the NASA Graduate Student Researchers Program.					
16. Abstract Transmitting large amounts of data efficiently among neighboring stars will vitally support any eventual contact with extrasolar intelligence, whether alien or human. Laser carriers are particularly suitable for high-quality, targeted links. Space laser transmitter systems designed by this work, based on both demonstrated and imminent advanced space technology, could achieve reliable data transfer rates as high as 1 kb/s to matched receivers as far away as 25 pc, a distance including over 700, approximately solar-type stars. The centerpiece of this demonstration study is a fleet of 13 automated spacecraft incorporating adaptive neural-net optical processing, active structures, nuclear electric power plants, annular momentum-control devices, and ion propulsion. Together the craft sustain, condition, modulate, and direct to stellar targets an infrared laser beam extracted from the natural mesospheric, solar-pumped, stimulated CO ₂ emission recently discovered at Venus. For a culture already supported by mature interplanetary industry, the cost of building planetary or high-power space laser systems for interstellar communication would be marginal, making such projects relevant for the next human century. Links using high-power lasers might support data transfer rates as high as optical frequencies could ever allow. A nanotechnological society such as we might become would inevitably use 10 ²⁰ b/yr transmission to promote its own evolutionary expansion out into the galaxy.					
17. Key Words (Suggested by Author(s)) extraterrestrial intelligence, SETI, CETI, planetary lasers, Mars, Venus, CO ₂ lasers, adaptive control, optical communication, advanced spacecraft design, nanotechnology, 21st century, large space structures, interstellar transport				18. Distribution Statement Unclassified - Unlimited Subject Category - 18	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of pages	
				22. Price	

National Aeronautics and
Space Administration

Washington, D.C.
20546

Official Business

Penalty for Private Use, \$300

Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451



NASA

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return
